TOWARDS THE DEVELOPMENT OF A NOVEL COLOURIMETRIC NUCLEIC ACID BIOSENSOR BASED ON PEPTIDE NUCLEIC ACID-FUNCTIONALISED POLYDIACETYLENE LIPOSOMES

Jennyfer Goujon

Submitted for the degree of Doctor of Philosophy

Heriot-Watt University
School of Engineering and Physical Sciences
September 2009

ABSTRACT

The aim of the research project described here was to develop a novel colourimetric nucleic acids biosensor based on polydiacetylene liposomes containing lipophilic peptide nucleic acids. Preliminary investigations in this area have shown that PNA-containing PDA liposomes can be constructed and that they are blue in colour as expected. However, their poor water solubility and the resulting precipitation made it necessary to synthesise and evaluate a second generation. In these, the PNA head-group is separated from the lipid tail by an amino diethylene glycol-type spacer molecule.

The first hydrophilic spacer synthesised was 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctanoic acid. Three methods based on the *O*-alkylation of mono-Boc, di-Boc and dibenzyl 5-amino-3-oxapentanol were devised. The Boc strategy afforded the corresponding ether in low yield of 20% while in the case of the dibenzyl approach a suitable methodology to effectively cleave the protecting groups from the amino function could not be found. Moreover, the model reaction between dibenzyl 8-amino-3,6-dioxaoctanoic acid and thyminyl PNA monomer afforded the corresponding conjugates in only 6%.

An alternative type of connection to link the spacer to the PNA headgroup was thus sought. A 1,4-disubstituted [1,2,3]-triazole moiety was obtained in 60% yield by reacting spacer, 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctan-1-azide and *N*-alkynyl PNA monomers bearing thymine, Cbz protected adenine and Cbz protected cytosine according to the conditions of the 'Click' reaction. These 'spacer-PNA monomer' intermediates were then functionalised at their *N*-terminus using 10,12-pentacosadiynoyl fluoride, to afford 'lipid-spacer-PNA monomer' models. The saturated conjugates were obtained by first preparing the 'lipid-spacer' intermediate, *N*-(8-azido-3,6-dioxaoctanyl)stearamide which was then connected to the *N*-alkynyl PNA monomers using the 'Click' reaction. Following this last strategy, a homothymine PNA dimer, prepared in solution phase, was also functionalised with the saturated lipid chain.

This second generation of lipid functionalised-PNA monomer models were incorporated into PDA-liposomes. Upon photo-polymerisation, the liposome solutions became blue. They were found to be stable in solution, at 4 ^{o}C over long period of time. Their colour changed to red upon environmental changes (i.e. pH and temperature).

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Dr Nicola Howarth for her constant guidance, encouragement and support over the last few years, and for her interest for the project. I also thank Dr Arno Kraft for his advice.

I would also like to express gratitude to Mrs Christina Graham for elemental analysis, Dr Alan Boyd for NMR spectroscopy, Dr Georgina Rosair for X-Ray diffractometry, the EPSRC Mass Spectrometry Service Centre at the University of Wales, Swansea for HRMS measurements and Heriot-Watt University for the funding allowed to support my work.

This task would not have been manageable without my friends and family who have supported me over the past few years. Thanks to Carla for her continued friendship and support over the past few years. Thanks also to Neil, Barbara, Ben, Atchma, Rob, Jiang, Mary and John who have made my time spent at Heriot-Watt University more enjoyable. A particular thanks to Lia and Koen my fellows and friends for their never-lasting good mood, supports and the good moments shared with them in the lab.

I would especially like to thank my parents as well as Eve and Julien for the constant support and my fiancé, Nicolas, for his much needed encouragement and patience throughout this PhD.

ACADEMIC REGISTRY
Research Thesis Submission

Name*:*			
School/PGI:			
Version: *(i.e. First, Resubmission, Final)*		Degree Sought (Award **and** Subject area)	

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1) the thesis embodies the results of my own work and has been composed by myself
2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted*.
4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate*:*		Date:	

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in Academic Registry

Received in the Academic Registry by *(name in capitals):*			
Method of submission *(Handed in to Academic Registry; posted through internal/external mail):*			
E-thesis Submitted ***(mandatory for final theses from January 2009)***			
Signature:		Date:	

CONTENTS

ABBREVIATIONS

A	Adenine
Ac	acetyl
AEEA	8-amino-3,6-dioxaoctanoic acid
ATP	adenosine triphosphate
B	nucleobase
Boc	*tert*-butoxycarbonyl
br	broad
C	Cytosine
Cbz	benzyloxycarbonyl
CR	colourimetric response
CTP	cytidine triphosphate
d	doublet
DABCYL	[4-((4-(dimethylamino)phenyl)azo)benzoic acid]
DADMDPA	3,3'-diamino-*N*-methyldipropylamine
DCC	N,N'-dicyclohexylcarbodiimide
DCM	dichloromethane
DhbtOH	3-hydroxy-1,2,3-benzotriazin-4(3H)-one
DIPEA	diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMPC	dimyristoylphosphatidylcholine
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
ds	double-stranded
DSC	Differential Scanning Calorimetry
dt	doublet of triplets
EDEA	[2,2'-(ethylenedioxy)*bis*(ethylamine)]
EI	electron impact
equiv.	equivalents
ESI	electron spray ionisation
Et	ethyl
EtOAc	ethyl acetate
Fmoc	9-fluorenylmethoxycarbonyl

G	Guanine
Glu	glutamic acid
GTP	guanosine triphosphate
h	hours
HA	hemaglutinin
HATU	2-(*1H*-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl-uronium hexafluoroborate
HBTU	2-(*1H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HCDA	heneicosadiynoic acid
HCl	hydrochloric acid
HCMV	human cytomegalovirus
HIV	human immunodeficiency virus
HRMS	High Resolution Mass Spectrum
Im	Imidazole
IR	infra-red
LB	Langmuir-Blodgett
LS	Langmuir-Schaefer
Lys	lysine
M	molar
m	multiplet
ma.	major rotational isomer
MB	Molecular Beacon
Me	methyl
min	minutes
min.	minor rotational isomer
mRNA	messenger ribonucleic acid
MsCl	methanesulfonyl chloride
NA	neuraminidase
NMR	Nuclear Magnetic Resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
OTS	octadecyltrichlorosilane
PBS	phosphate buffered saline
PCDA	pentacosadiynoic acid

pcPNA	pseudo-complementary PNA
PCR	polymerase chain reaction
Pd/C	palladium on charcoal
PDA	polydiacetylene
PEG	poly (ethylene glycol)
Ph	phenyl
pK_a	Ionisation constant
PLA_2	phospholipase A_2
PLC	phospholipase C
PLD	phospholipase D
PNA	peptide nucleic acid or polyamide nucleic acid
PNAA	peptide nucleic acid amphiphile
ppm	part per million
PyBrop	bromotris(pyrrolidino)phosphonium hexafluorophosphate
q	quartet
QCM	Quartz Crystal Microbalance
Rf	retention factor
RNA	ribonucleic acid
ROE	rotating Overhauser effect
ROESY	rotating frame Overhauser effect spectroscopy
rRNA	ribosomal ribonucleic acid
rt	room temperature
s	singlet
SAM	self-assembled monolayer
SPR	Surface Plasmon Resonance
ss	single-stranded
ST-DNA	salmon testes DNA
T	Thymine
t	triplet
TAMRA	tetramethyl-6-carboxyrhodamine
TBAF	tetrabutylammonium fluoride
TBAI	tetrabutylammonium iodide
TBTU	2-(*1H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetra-fluoroborate
TCDA	tricosadiynoic acid

TEM	Transmission Electron Microscopy
tert	tertiary
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	Thin Layer Chromatography
Tm	melting temperature
TMS	trimethylsilyl
tRNA	transfer ribonucleic acid
U	Uracil
UTP	uridine triphosphate
UV	Ultra-Violet
v/v	volume/volume

Note on abbreviation of PNA oligomers:

The nomenclature introduced by Nielsen *et al.*[1] to denote PNAs has been used in this thesis to report the novel PNAs synthesised by Murray *et al.*.[2] For the oligomer H-TCAG-Lys-NH_2, H denotes the free *N*-terminus amino group; T, C, A and G the T-, C-, A- and G-acetyl *N*-(2-aminoethyl)glycine units, respectively; and Lys-NH_2 indicates a *C*-terminus lysine amide. Additionally, $CH_3(CH_2)_{16}CO$, $CH_3(CH_2)_{11}C_4(CH_2)_8CO$ and Ac has been employed to denote the acyl group attached to either the *N*-terminus of the PNA oligomer or lysine side chain.

CHAPTER 1: INTRODUCTION

1 INTRODUCTION

The development of nucleic acid biosensors capable to quickly provide precise information on particular sequence of genes is an area of bio-technological research that has received particular attention in recent years. Such an interest has been followed by the research group at Heriot-Watt University since it has embarked into a research programme to develop a novel type of colourimetric nucleic acid biosensor. This thesis describes the progress made towards the development of lipid-functionalised peptide nucleic acids (PNAs). It is hoped that such materials can be used as nucleic acid biosensors.

The following introduction is a brief summary of recent research into the areas covered by the work. It includes a short introduction to nucleic acid structure and function, as well as a short summary on the type of nucleic acid biosensors that have been recently developed and that are currently available for research. PNA structures and uses are also described along with polydiacetylenes. In the latter case, the focus is made on their chromic properties and how these have been extrapolated for the development of a variety of colourimetric biosensors.

Subsequently, the attention of the thesis will turn to how the research group have brought together these fields in an attempt to construct nucleic acid biosensors.

1.1 NUCLEIC ACIDS: STRUCTURES AND FUNCTION

1.1.1 Nucleic acid structures

There are two classes of nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). These were first discovered in 1869 by Friedrich Miescher.[3] In most living organisms, the genetic information is carried by genes, segments of DNA, contained within macromolecular structures called chromosomes. The exceptions are for some viruses which transport their hereditary information as RNA rather than DNA. Aside from long-term storage of the genetic information, DNA plays two roles: it directs its own replication during cell division and it directs the production of complementary molecules of RNA during the process of transcription. RNA may be associated with proteins through formation of complexes called ribonucleoproteins. All these RNA

molecules are required for the biosynthesis of proteins through the process of translation.

Nucleic acids are linear polymers of monomeric units called nucleotides. Each nucleotide consists of three components: a nitrogenous heterocyclic base, a pentose sugar and a phosphate group. The phosphate unit bridges the 3′- and 5′- positions of successive sugar residues. The monosaccharide unit of RNA is β-D-ribofuranose (usually referred to as ribose) whereas in DNA, it is β-2-deoxy-D-ribofuranose (or deoxyribose), where the 2'-hydroxyl group of ribose has been replaced by an hydrogen. The nitrogenous heterocyclic bases of nucleic acids, or nucleobases, are separated into two families: (i) the pyrimidines of thymine (T), cytosine (C) and uracil (U); and (ii) the purines of adenine (A) and guanine (G). These nucleobases are attached to the sugar moiety through a linkage formed either between *C*-1′ of the sugar and *N*-9 of the purines or through *C*-1′ of the sugar and *N*-1 of the pyrimidines. Adenine, guanine and cytosine are shared by DNA and RNA, whereas thymine is only found in DNA; it is exchanged for uracil in RNA (*Figure 1*). The linear polymer of nucleotide subunits is called oligonucleotide chain. These chains have directionality; the terminus bearing the free 5′-hydroxyl on the sugar is referred as the 5′-end whereas the one terminating with a 3′-phosphate moiety is called the 3′-end.

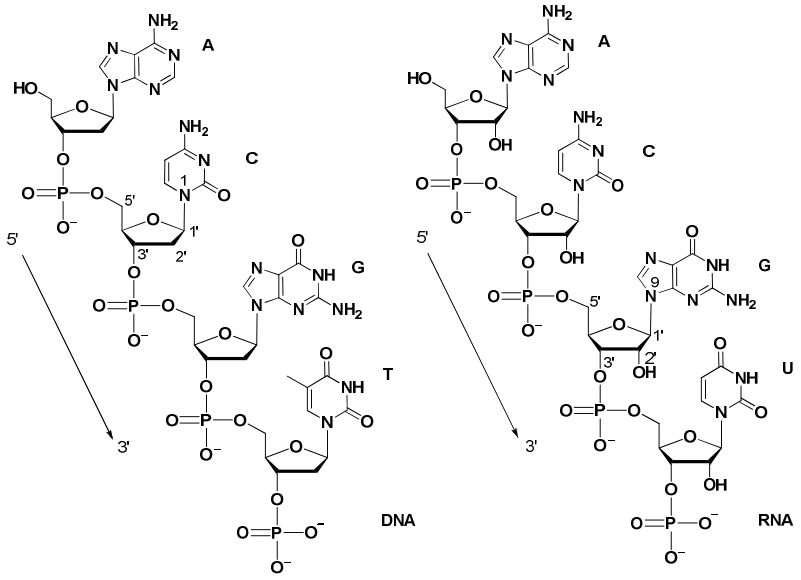


Figure 1 – Structures of DNA and RNA.

The structure of DNA and its key role in holding all the information required for cell reproduction and function were first described by James Watson and Francis Crick in 1953.[4, 5] This discovery is often associated with the birth of modern molecular biology. Watson and Crick found that DNA is composed of two oligonucleotide chains coiled around each other to form a right handed double helix (*Figure 2*). These strands are held together through the formation of specific hydrogen bonds between the nucleobases which reside in the centre of the helix. The sugar phosphate backbones coil about the periphery of the helix, in such a way as to minimise the repulsion between the negative charges on the phosphate units. The two strands line up in an anti-parallel orientation (i.e. in opposite direction) and about a common central axis to form a double helix of diameter *ca.* 2 nm.

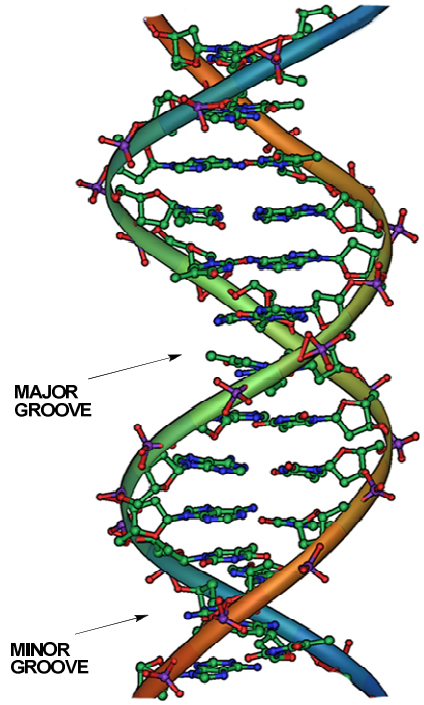


Figure 2 – The structure of a part of DNA double-helix

Each base is hydrogen bonded, in a phenomenon known as complementary base pairing, to a specific base on the opposite strand. This fact contributes to the fidelity of DNA replication. Double-helical nucleic acid molecules contain two grooves, called the major groove and the minor groove (*Figure 2*). These grooves arise because the glycosidic bonds of a base pair are not diametrically opposite to each other (*Figure 3*).

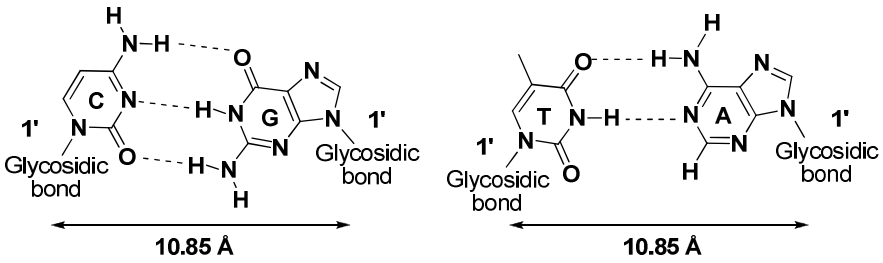


Figure 3 – Base pairing in DNA. Hydrogen-bonding shown in dotted lines.

There are only two types of base pairs found in DNA, A-T and G-C, namely the Watson-Crick base pairs (*Figure 3*). The distance between the two C-1' atoms of the sugar phosphate backbone is the same in the two base pairs so that they can be interchanged within the double helix without distortion of the structure (*Figure 3*). This structure allows for only sequence of bases on one strand if the opposite strand has the complementary base sequence, any other combination of bases would result into a considerable reorientation of the polymer chain. This complementarity let Waston and Crick[4, 5] to suggest that hereditary information is encoded in the base sequence on either strand.

The structure of DNA described by Watson and Crick is now known as B-DNA but double-stranded (ds) DNA can also assume two others conformations, namely A-DNA and Z-DNA.

	A-DNA	**B-DNA**	**Z-DNA**
Shape	Broadest	Intermediate	Narrowest
Rise per base pair	2.3 Å	3.4 Å	3.8 Å
Helix diameter	25.5 Å	23.7 Å	18.4 Å
Screw sense	Right-handed	Right-handed	Left-handed
Base pairs per turn of helix	11	10.4	12
Major groove	Narrow and very deep	Wide and quite deep	Flat
Minor groove	Very broad and shallow	Narrow and quite deep	Very narrow and deep

Table 1 – Comparison of the different forms of DNA

In B-DNA, the major groove is deeper (8.5 Å instead of 7.5 Å) and wider (12 Å instead of 6 Å) than the minor groove. The larger size of the major grove in B-DNA makes it more accessible for interactions with proteins that recognize specific DNA sequences. Under physiological conditions, most DNA is in the B form. When the relative humidity

is reduced to less than *ca.* 75%, then A-DNA is found. A-DNA is a right-handed double helix like B-DNA, but it is wider and shorter and its bases pairs are titled (by 19^{o}) rather than perpendicular to the helix. On the other hand, Z-DNA is a left-handed helix in which the backbone phosphate zigzag. This form is adopted by short oligonucleotides that have a sequence of alternating purines and pyrimidines. High salt concentrations are required to minimize electrostatic repulsion between the backbone phosphates, which are closer to each other than in A- and B-DNA. The properties of A-, B- and Z-DNA are summarised in *Table 1*.[6]

The second type of nucleic acids present in nature is RNA and is as important as DNA. Unlike DNA, an RNA molecule usually exists as a single strand, although significant segments may contain self-complementary sequences that allow parts of the RNA to fold and pair with itself to form double helices. Cells contain several types of RNA, the three majors are listed below:

- Messenger RNA (mRNA) is the template that carries information from DNA and is used in the synthesis of proteins (see later Sections 1.1.3 and 1.1.4). The coding sequence of the mRNA determines the sequence of amino acid in the protein.[6]
- Transfer RNA (tRNA) carries amino acids in an activated form to the ribosome for peptide-bond formation, the sequence is dictated by the mRNA template. It consists of *ca.* 75 nucleotides and there is at least one type of tRNA for each of the twenty amino acids.[6]
- Ribosomal RNA (rRNA) is the major component of ribosomes and plays both a catalytic and a structural role in the synthesis of proteins. The ribosome binds mRNA and carries out protein synthesis. It is the most abundant type of RNA in eukaryotic cytoplasm, *ca.* 80%.[6]

1.1.2 DNA Replication

The process of DNA replication is semi-conservative, as each strand of the polynucleotide acts as a template for the reproduction of the Watson-Crick complementary strand through base-pairing interactions. The DNA molecule unwinds to form a replication fork and the two daughter strands are then synthesized by enzymes called DNA polymerases (*Figure 4*). The strand running from 5′ to 3′ is called leading strand and is synthesized continuously by the addition of nucleotides donated by a nucleoside triphosphate to the 3′-OH group of a base-paired polynucleotide. The other strand is called the

lagging strand and is synthesized in small segments known as Okasaki fragments. These are subsequently joined together by the enzyme DNA ligase to give a complete strand. The process of replication gives two identical dsDNA molecules, each having one strand from the parent molecule and a new one.

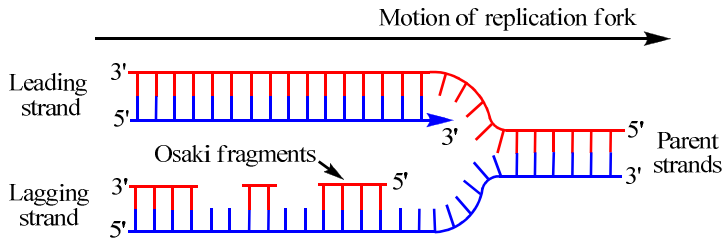


Figure 4 – Schematic representation of the process of DNA replication

1.1.3 Transcription

Since DNA is confined in the nucleus in the eukaryotic cell and protein synthesis occurs in the cytoplasm, processes that allow the transfer of information contained on DNA are required. Transcription is the process of making RNA from a DNA template. This starts with the unwinding of the DNA double helix by the enzyme helicase, in order to expose the template strand running from 5′ to 3′. Then several transcription factors bind to the DNA, upstream to the portion of the DNA used to produce RNA (i.e. transcription unit) and help the successful binding of RNA polymerase. The RNA polymerase then adds ribonucleoside triphosphate monomers (GTP, CTP, ATP and UTP) that are coupled to their complementary Watson-Crick base on the template strand. The transcription resulting in a growing sequence of RNA continues until the RNA polymerase passes through a terminator sequence. This triggers RNA polymerase to release the DNA and ends transcription.

1.1.4 Translation

Translation is the process of synthesising a protein from an mRNA template. The process begins with the small unit of ribosome binding onto the 5′ end of mRNA and moving to the initiation site. One of the key molecules involved in the process is tRNA. This is a short sequence of RNA made of 60-95 nucleotides, carrying an anti-codon complementary to mRNA codons and the corresponding amino acid. Complementary tRNA binds to the mRNA start codon and this event triggers the association of the large sub-

unit of ribosome to the small one, both bound to mRNA. This association creates the P-site (peptidyle) and the A-site (amino acyle); the first tRNA occupies the P-site, while the second tRNA coming next, occupies the A-site (and is complementary to the second mRNA codon). The first amino acid is then transferred to the A-site amino acid, the first tRNA exits the P-site and the ribosome moves along the mRNA sequence, this is the process of elongation. When a stop codon is encountered in the A-site, a release factor enters and binds, the translation is terminated. This triggers the ribosome to dissociate and the release of the newly formed peptide chain. This finally undergoes post-translational modifications in order to become a mature protein.

1.1.5 Mutations of DNA

The process of DNA replication (Section 1.1.2) sometimes leads to nucleotides copying errors called mutations. These mutations can also be caused by viruses, exposure to chemicals, UV- and ionizing irradiations. Gene mutations have varying effects depending on where they occur; they may alter the function of the produced protein from translation of this gene and lead to serious health conditions. Mutations are classified in two families depending on the effect on structure: the small-scale mutations are those affecting a small gene in one or few nucleotides, while the large-scale mutations affect the chromosomal structures and are therefore more drastic.

Here, we are interested in small-scale mutations; these are again subdivided into three classes: insertions, deletions and point mutations. Insertions add one or more nucleosides into the DNA while deletions remove one or more nucleosides from the DNA. Both cause significant alterations to the gene product as the insertion or the deletion of a number of nucleotides that is not divisible by three can cause disruption in the grouping of codons, resulting in completely different translation product from the original (see Section 1.1.4). If such insertion or deletion occurs early in the sequence, the product obtained will be more altered. On the other hand, point mutations (also called single nucleotide polymorphisms or SNPs) are the exchange of one nucleotide for another one. Two exchanges are possible: the exchange between two nucleobases of the same nature (pyrimidine for pyrimidine or purine for purine), which is the most common point mutation and is called transitions; the other possibility, less common, is called transvertions and is the exchange of a purine for a pyrimidine or *vice-versa*. Point mutations can arise anywhere on DNA and those occurring within the protein coding region of a gene can

have dramatic effects. Mutation can be neutral when it results in the coding of a different but chemically similar amino acid. The effects are usually negligible and this is therefore very similar to silent mutation. Silent mutation leads to the coding of the same amino acids from a different codon (e.g. codon AU**U** exchanges for codon AU**C** still encode leucine) and has no discernible effects on the protein produced. Nonsense mutation results in a premature stop codon, leading to a truncated and often non functional protein product. Finally, missense mutation results in the coding of a different amino acid, leading to a non-functional protein. Such point mutations are responsible for human diseases such as sickle-cell disease and are found in *p53*. Protein 53 is a transcription factor (i.e. a protein that binds to specific DNA sequences controlling the transcription process of this portion) which is encoded by the *p53* gene. It is part of the regulatory cell cycle and acts as a tumour suppressor. On this gene, the replacement of guanine by adenine at base 524 leads to an altered *p53* protein with an arginine being replaced by a histidine amino acid. This mutation is found in *ca.* 50% of human tumours and has been associated to aggressive growth of several types of cancer.[7] Sequence-specific analysis of the *p53* gene can therefore become very useful to assist monitoring of cancer progress.

Therefore, due to the impact that such small changes can have on human proteome, it is evident that the ability to detect these mutations both efficiently and rapidly is of importance for health care.

1.2 NUCLEIC ACID-DETECTION TOOLS

1.2.1 Background

From Section 1.1.5, the development of sensitive, rapid and simple diagnostic tools, such as biosensors, therefore offer great promise for DNA diagnostics.

Biosensors are small devices employing biological recognition properties for a selective bioassay. They rely on the intimate coupling of a biological recognition element with a physical transducer. The recognition element (also referred to as recognition layer) enables selective detection of the analyte, while the transducer converts the recognition event into a useful electronic signal proportional to the concentration of the target analyte (*Figure 5*). The signal obtained can take a variety of forms such as a current, poten-

tial, temperature change, absorption of light, or mass increase obtained through electrochemical, optical, or piezoelectric means. The signal may be further amplified, processed, or stored for later analysis.[8]

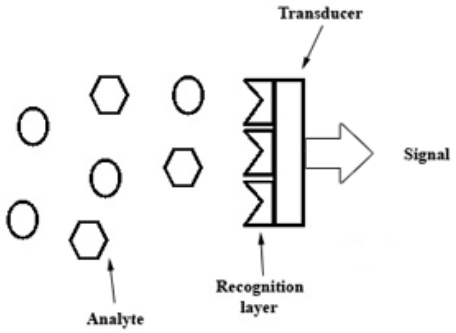


Figure 5 – Principle for the detection of an analyte using a biosensor.

DNA diagnostic tools have been developed for the detection and analysis of genes. The basis of these tools is the base pairing feature of DNA. They rely on the immobilisation of a small ssDNA probe (i.e. 25 to 40 mer) of known sequence onto the surface of the transducer. When the biosensor is exposed to a sample solution containing the target sequence, the target ssDNA molecules bearing the complementary sequence hybridise to the ssDNA probe at the surface of the transducer. This causes the transducer to send out the signal informing of the successful hybridisation event (*Figure 6*). A perfect match in the target sequence produces a very stable dsDNA complex, while the presence of one or more base mismatches decreases the stability, causing modifications in the signal.

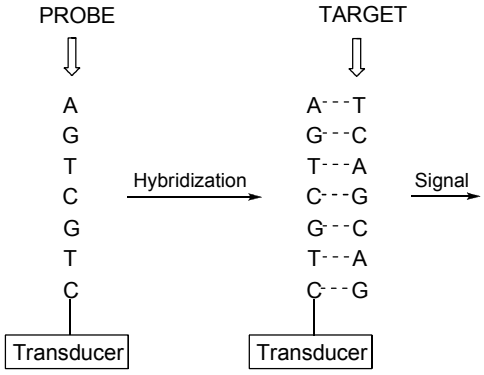


Figure 6 – General DNA biosensor design.

Two aspects are essentials when developing hybridisation biosensors: the sensitivity and the selectivity. It is mandatory that the tool is able to detect DNA at low concentration and to detect point mutations. Therefore, one requirement to obtain such a biosensor is

to ensure that the probe sequences are properly orientated (i.e. vertical) to allow maximum accessibility from the target molecule. The immobilisation step for the DNA probe must, therefore, be controlled. Different methods have been reported in the literature and the most common includes electrostatic adsorption,[9] covalent immobilisation,[10] and avidin (or strepavidin)-biotin interactions.[11, 12]

Electrostatic adsorption has the advantage that it is fairly simple as it does not require any nucleic acid modification but has its own drawbacks. The principle of the method is based on ionic interactions occurring between negatively charged phosphate backbone of the DNA probe and positive charges covering the surface. The problem with this method is that the interactions are not specific; therefore, the strong affinity of DNA for the surface of the transducer results in the probe being fixed on the surface in multiple points and not just at one end as schematically represented *Figure 6*. This leads to low hybridisation efficiencies as the configurational freedom of the probe is restricted. Another drawback is that the target DNA will also have affinity for the transducer's surface, leading to non-specific adsorption of DNA.

As an alternative to adsorption, one end of the probe strand can be modified with a reactive group that can interact with the transducer surface leading to the formation of a covalent bond. An example of this is to the modification of the DNA probe at one of its end with a mercaptoalkyl linker (i.e. thiolated-DNA). This one point attachment method uses the strong affinity of thiol groups for gold surfaces, ensuring a strong bond between the DNA probe and the transducer surface (*Equation 1*).[13]

$$\text{R-SH} + \text{Au} \longrightarrow \text{R-S-Au} + e^- + H^+$$

Equation 1

Thiolated-DNA self-assembles on the gold surface but the use of mercaptoalkyl linkers alone was not sufficient to ensure DNA proper orientation. In fact, a low hybridisation efficiency with target (between 0 and 5 %) is obtained, probably because of the affinity of the DNA bases for gold surfaces.[14] This efficiency was greatly improved when the DNA-modified surface was incubated in a solution of mercaptohexanol used as a diluent layer (if the alkyl linker used was mercaptohexane). The affinity of the thiol moiety of the mercaptohexanol diluents for the gold surface allowed the DNA probe to be

raised off the surface. In addition, the negative dipole of the alcohol on the distal end repelled the negatively-charged backbone of DNA that resulted in a vertical position (*Figure 7*).[15]

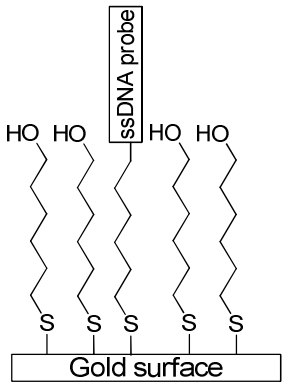


Figure 7

Finally, the last immobilisation method commonly used involved the formation of non-covalent avidin (or streptavidin)-biotin complex. Biotin (*Figure 8*) is a small molecule that binds with high affinity to the proteins avidin and strepavidin.[11]

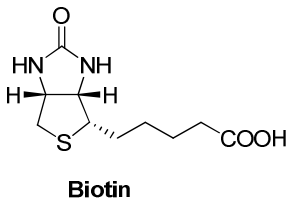


Figure 8

This procedure is frequently used in the preparation of DNA biosensors as the biotin-streptavidin complex is one of the strongest non-covalent interactions known in nature. In a general procedure, the avidin is first deposited onto the surface of a graphite electrode, then the modified electrode is dipped in a solution containing the DNA probe which had been first biotinylated at its 5′-end, resulting therefore in the immobilisation of the probes. Finally, the electrode-modified with DNA probe is subsequently incubated with solutions containing the targets for analysis.[12]

The surface coverage is another important feature that needs to be controlled in order to prepare a sensitive biosensor. One could think that the larger the number of probes presents at the surface of the transducer, the greater the signal response and therefore, the

more sensitive the device. But if the space between probes on the surface is not sufficient repulsion between the incoming DNA targets may hamper hybridisation which will therefore lead to low detection efficiency. An ideal surface coverage of *ca.* 5×10^{12} molecules/cm^2 had been proposed by Peterlinz and Georgiadis.[16]

In the following three subsections, three types of transducers (i.e. electrochemical, optical, and piezoelectric) currently available on the market will be briefly introduced and compared.

1.2.2 Optical transducers

Several different optical transducers are currently available for use in the development of biosensors, including fluorescence, colourimetry, surface plasmon resonance (SPR), chemiluminescence, interferometry and surface-enhanced Raman scattering (SERS) spectroscopy. All of these transducers are very sensitive and offer great platforms for the detection of nucleic acids.

The most common DNA optical biosensors rely on a fibre optic to transduce the emission signal of a fluorescence tag. Krull *et al.*[17] developed the first biosensor of this type which used ethidium bromide (EB) as the fluorescent label for recognising dsDNA (*Figure 9*).

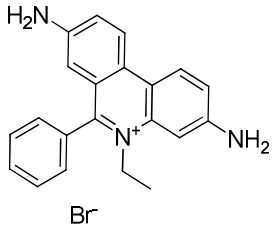


Figure 9 – Chemical structure of ethidium bromide

The fluorescent ethidium cation (3,8-diamino-6-phenyl-5-ethyl-phenanthridium) is well known to strongly associate with dsDNA by intercalating between base pair (π-stacking interactions).[17, 18] It was this strong interaction of EB for dsDNA which was explored in Krull's biosensor. In their approach, the target sample of dsDNA was first exposed to an EB solution. Subsequently, any remaining unbound EB was removed by washing. The fluorescence intensity of EB was recorded by UV-visible spectroscopy. Since the fluo-

rescence intensity is directly proportional to the amount of EB bound to DNA, the amount of dsDNA in the original sample could be determined. Today, the original mutagenic EB employed by Krull *et al.*[17] has been largely replaced by Synergy Brands, Inc. (SYBR®) stains; these have become very popular due to their sensitivity and ease-of-use. SYBR green I (*Figure 10*), SYBR green II and SYBR gold are by far the best high-sensitivity reagents commercially available.[19]

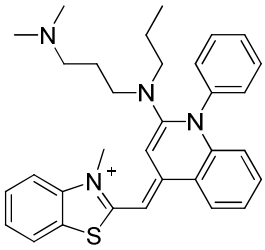


Figure 10 – Chemical structure of SYBR green I

The quantum yields for these new stains are many folds superior to EB. *Table 2* summarises the excitation and fluorescent emission characteristics of these stains.

Stain	Excitation max	Fluorescent emission max
SYBR Gold	~ 495 and 300 nm	~ 537 nm
SYBR Green I	~ 497, 380 and 290 nm	~ 520 nm
SYBR Green II	~ 497 and 254 nm	~ 520 nm

Table 2 – Comparison of excitation and emission maxima.

SYBR gold has a high quantum yield (~0.6) upon binding to ds and ssDNA or to RNA. SYBR Green II is highly sensitive for detecting RNA or ssDNA with a quantum yield of (~0.54) when bound to RNA. Its quantum yield to dsDNA is lower (~0.36) making it the stain of choice for RNA detection. Finally, the quantum yield of dsDNA/SYBR Green I complex is *ca.* 0.8 and it is much lower for ssDNA and RNA, so this stain can be used for the detection of dsDNA in complex mixture. These three staining reagents offer a great fluorescence enhancement upon binding to nucleic acids. Although, SYBRs dyes are less toxic than EB, they are still prepared as dimethyl sulfoxide (DMSO) solutions; thus, there is a risk of their skin penetration for the operator. With the introduction of biosensor technology, the use of staining reagent has been replaced by fluorescent labels that are incorporated into the DNA probe using either enzymatic incorporation (RNA or DNA polymerase) or the direct labelling strategy. For example, KREACHER Biotechnology BV have developed and patented a direct labelling method

using a platinum compound that connects the label in the *N*-7 position of guanine bases. Alternatively, labels can be added to the 3′-terminus of the probe by using the enzyme terminal deoxynucleotidyl transferase (TdT) in combination with labeled uridine triphosphates (UTP), deoxyuridine trisphosphates (dUTP) or deoxycytidine triphosphate (dCTP) nucleotides. Molecular Probes, a subsidiary of Invitrogen Ltd., has developed a series of fluorescent dyes that span the whole visible spectrum (*Figure 11*) with a superior brightness than common dyes, these are called Alexa® Fluor. Alexa® Fluor dyes are synthesised *via* sulfonation of coumarin, rhodamine, xanthenes and cyanine dyes. These novel fluorophore are more photostable allowing longer recording time of the image obtained. Sulfonation leads to negatively charged molecules that are therefore more soluble in water. These dyes are also less sensitive to pH than common dyes. But they are also more expensive and patented by Invitrogen Ltd.

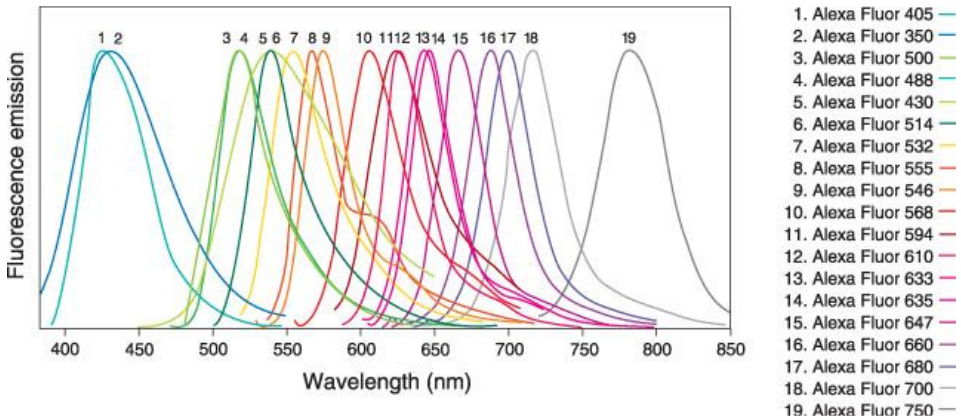


Figure 11 – Diagram representing the fluorescence emission of Alexa® fluor dyes.

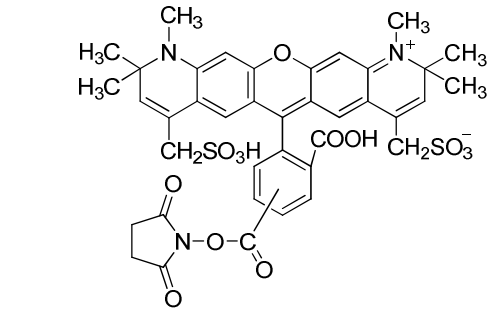
Table 3 summarises the excitation and emission maxima as well as colour obtained.

Fluorophore	Ex*	Em*	Colour	Similar dyes
Alexa Fluor 350	346	442	Blue	AMCA
Alexa Fluor 488	490	520	Green	Fluorescein
Alexa Fluor 532	534	550	Yellow	Rhodamine 6G
Alexa Fluor 546	561	570	Orange	Cy3 dye
Alexa Fluor 568	575	600	Red	Lissamine™ rhodamine B dye
Alexa Fluor 594	590	615	Red	Texas Red dye
Alexa Fluor 647	650	670	Far-red	Cy5 dye

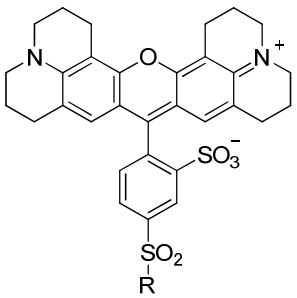
* Excitation (Ex) and Emission (Em) maxima, in nm

Table 3

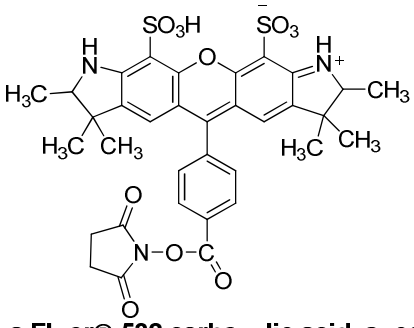
The structures of some Alexa® Fluor labels and others dyes commercially available are represented in *Figure 12*.



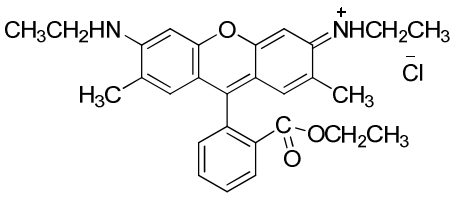
Alexa Fluor® 594 carboxylic acid, succinimidyl ester



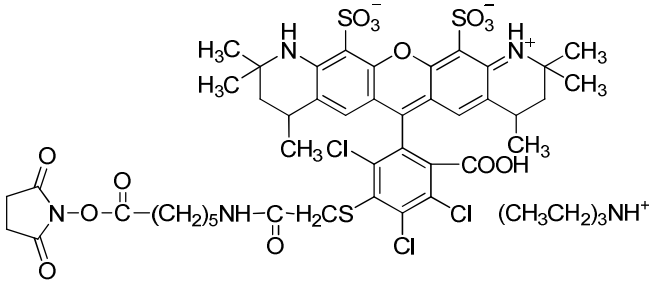
Texas Red®



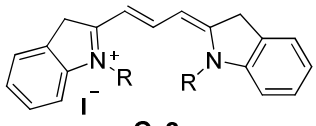
Alexa Fluor® 532 carboxylic acid, succinimidyl ester



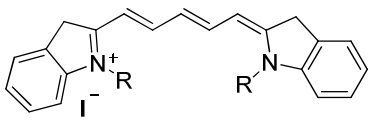
Rhodamine 6G chloride



Alexa Fluor® 546 carboxylic acid, succinimidyl ester



Cy3



Cy5

Figure 12 – Structures of commonly used fluorescents dyes.

The spectral diversity offered by these fluorescent labels (*Table 3*) gives researchers the possibility to carry out multicolour fluorescence hybridisation strategies, commonly carried out on microarrays. It is possible to distinguish four to five different fluorescent signals in a single sample using their excitation and emission properties alone along with the appropriate optical filters. The use of microarray minimises the artefactual variation in signals. These fluorescence-based optical biosensors offer increased sensitivity with nanomolar concentration detection limits. But this method has disadvantage of the requirement for the use of highly precise and expensive instrumentation [such as microscopy-coupled charge-coupled device (CCD) cameras, fluorescent scanner]. So-

phisticated numerical algorithms to interpret the data are also mandatory, which therefore limit the method to research laboratories principally. Another problem commonly reported is the inconsistencies in yields of target DNA synthesis and labelling, as well as non-uniform rates of fluorophore photobleaching that can make the accuracy of the assay lower than what is required for gene analysis.[20]

An alternative strategy which has been employed to detect DNA by fluorescence, involves the use of molecular beacons (MBs). MBs are ssDNA probes that possess a stem-and-loop structure.[21, 22] The loop portion is complementary to the DNA target. A fluorophore (e.g. 6-carboxytetramethylrhodamine, TAMRA) and a quencher (e.g. (4-(4-(dimethylamino)-phenyl)azo)benzoic acid DABCYL) are linked to the two ends of the stem (*Figure 13*).

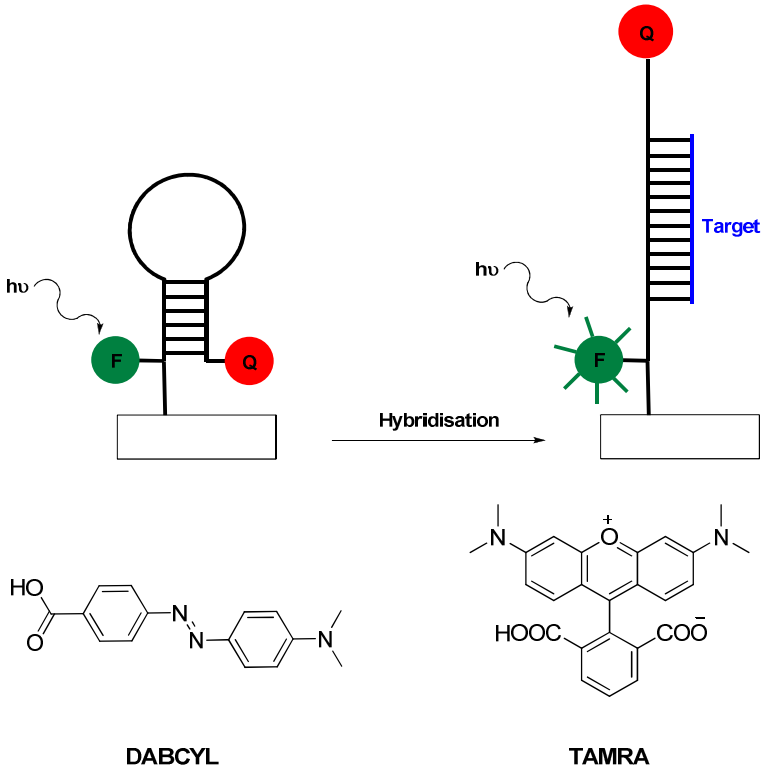


Figure 13 – Immobilised molecular beacon's strategy and structures of a fluorophore and a quencher.

In its native state, the probe is a hairpin; in this form, the fluorophore lies close to the quencher and thus its fluorescence is quenched by energy transfer (*Figure 13*). However, in the presence of the DNA target, the MB undergoes conformational reorganization to allow the loop to hybridise to the target. As a result the probe adopts a single strand structure which separates the fluorophore and the quencher. Thus, this reorganisation leads to emission of an intense fluorescent signal. Biosensors based on MB tech-

nology have been found to be more sequence selective than devices using linear probe. The detection limit is *ca.* 10^{-9} M.[23, 24]

Surface Plasmon Resonance (SPR) is another commonly used optical technique that has been extensively developed in recent years in nucleic acid biosensors.[25] SPR can be utilised to monitor refractive index changes at a surface. When light is incident on a prism side at a particular angle, called the resonance angle, the intensity of the reflected light is at its minimum. When biomolecules are present on the surface, this variation angle is very sensitive and the changes in reflectivity give a signal proportional to the amount of the target bound to the surface.[13] Nucleic acid biosensors based on SPR are typically prepared by immobilising ssDNA probes onto the metal surface (e.g. gold); when the target DNA subsequently binds to the probes, the mass increases leading to an increase of the refractive index.[26] SPR has the advantage of being able to detect nucleic acids without labels and allows the measurement of the kinetics of biomolecular interactions in real time with high degree of sensitivity (detection limit *ca.* 10^{-9} M).[13] Furthermore, this method is suitable in array-based formats. However, the employment of SPR is hindered by a few disadvantages which include the cost and complexity of the technology required, making this technique more suitable for research applications.[27, 28] In addition, it has been reported that the sensitivity of the SPR technique is altered in the case of short-chain DNA molecules (15 mer) or highly diluted target solutions: in such cases, the resonance shifts are too small to facilitate detection. However, Liebermann *et al.*[29, 30] have reported a method that combined SPR and fluorescence detection. In their approach, the resonant excitation of an evanescent surface plasmon is used to excite a fluorophore that is chemically attached to the nucleic acid target. Upon binding to the probe immobilised at the surface of a metal, this chromophore reaches the strong optical fields that are obtained at the surface plasmon resonance giving rise to significant enahancement.[30]

1.2.3 Electrochemical transducers

In electrochemical DNA biosensors, the transducer is an electrode onto which the DNA probe has been immobilised. This method involves monitoring the current response produced upon exposure of the electrode to the target ssDNA as a result of the hybridisation event, under controlled potential conditions.[31] The probe-coated electrode is immersed into a solution containing the target DNA sequence to be tested. The target

DNA, which contains a sequence that matches that of the immobilised probe, binds to the probe and a DNA duplex forms at the surface of the electrode.[32] The formation of this duplex is then detected *via* the increased current signal of a redox indicator. Two approaches are then possible: these are referred to as 'label' method and the 'label-free' method.

'Labelled' methods use redox active molecules, that are present in the solution (e.g. enzyme,[33] intercalative redox substance,[34, 35] an interactive electroactive substance[36]), which preferentially interact with dsDNA, for example $Co(phen)_3^{3+}$ (*Figure 14*).

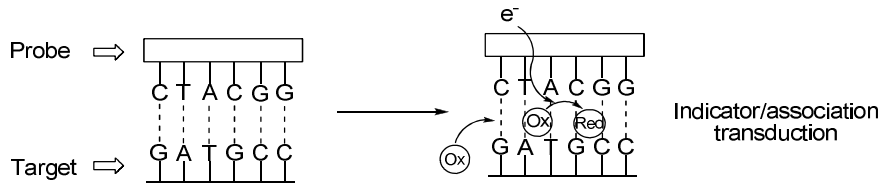


Figure 14 – Diagram of an electrochemical DNA biosensor based on redox indicators.

This area of research was pioneered by Mikkelsen's group[36] which first demonstrated the possibility of using the minor groove binding redox indicator, $Co(Phen)_3^{3+}$, for the detection of the cystic fibrosis ΔF508 deletion sequence.[36] In their work, the ssDNA probe was immobilised at the surface of the electrode and hybridisation to the target caused an increased in the surface concentration of electrostatically bound $Co(Phen)_3^{3+}$ resulting in a higher negative charge density at the surface. The readout was then provided through characteristic redox reaction of the cobalt.

Brossier and Limoges[33] have reported the development of an electrochemical biosensor which uses an enzyme as the label and used it for the detection of a gene sequence related to the human cytomegalovirus (HCMV). The probe of this biosensor was prepared from amplified HCMV ssDNA that has been adsorbed onto the surface of a screen-printed carbon electrode. The target sample was a biotinylated-modified ssDNA strand complementary to a portion of the probe. The hybridisation event was detected *via* streptavidin which was conjugated to horseradish peroxidase (HRP). The peroxidase label was, subsequently, indirectly quantified by measuring the amount of the chromophore and electroactive product 2,2'-diaminoazobenzene (**A** – *Figure 15*) generated from *o*-phenylenediamine. The intensity of the differential pulse voltammetric peak currents is related to the number of HCMV-amplified DNA molecules present in the sam-

ple. The detection limit obtained with this method was 0.6×10^{-15} mol / L (i.e. 230,000-fold more sensitive than gel electrophoresis).[33]

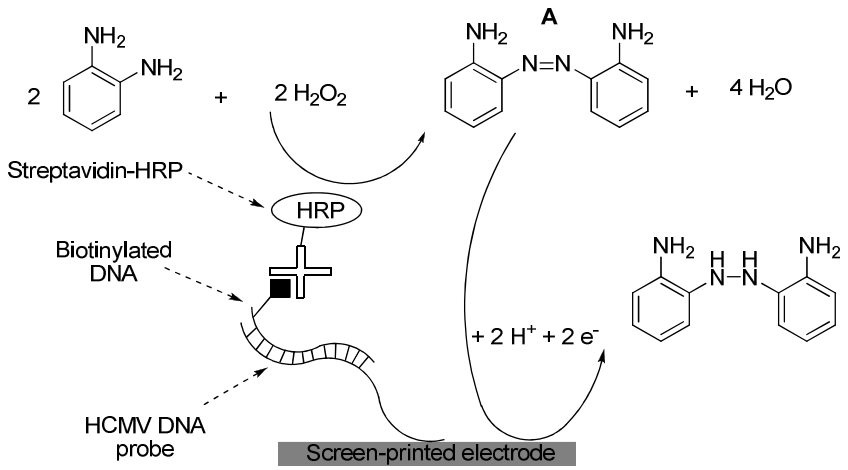


Figure 15 – DNA hybridisation detection using an enzyme.

New electroactive indicators, offering better distinction between ss and dsDNA have been reported for attaining higher sensitivity.[13, 32] For example, Takenaka *et al.*[35] have reported a detection limit of 10^{-21} mol when using the threading intercalator ferrocenyl naphthalene diimide (FND) that binds to the DNA duplex more tightly than usual intercalators and displays only negligible affinity for ssDNA probe (*Figure 16*).

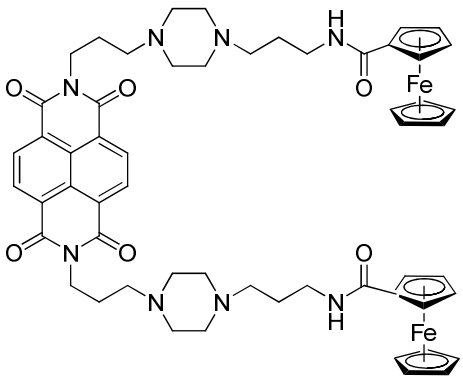


Figure 16 – Chemical structure of the FND.

As mentioned earlier 'label-free' electrochemical approaches are also possible, they rely on either the change to the electrical characteristics of the DNA-modified interface upon hybridisation or on the natural electroactivity of DNA.[32, 37] Such protocols greatly simplify the sensing operation and produce an instantaneous readout.[13] These assemblies operate by monitoring the changes in the intrinsic redox activity of the probe DNA induced during the hybridisation event. Of the four nucleic acids bases present, guanine is

the most easily oxidized and therefore the oxidation of this nucleobase can be monitored in indicator-free hybridisation detectors.[32] This method cannot be used for the detection of targets containing guanine bases. Wang *et al.*[37] have described a way to overcome this limitation, by using inosine-substituted (guanine-free) probes. Inosine preferentially forms base pair with the target cytosine residue and its oxidation signal is well separated from the guanine response. This allows direct detection of the DNA hybridisation through the appearance of the target guanine oxidation signal.

In summary, electrochemically based DNA hybridisation biosensors offer many advantages, such as high sensitivity, selectivity, low cost and the possibility for miniaturization, without the use of expensive instruments.

1.2.4 Mass-sensitive devices

The last transducer commonly used for the detection of DNA hybridisation uses a quartz-crystal microbalance (QCM). These sensors are made of a thin quartz disc sandwiched between a pair of electrodes. Quartz is a piezoelectric material meaning that it deforms when an electric field is applied across the electrode. The quartz crystal has a resonance frequency which is dependent on the total oscillating mass. This frequency increases with an increase in material mass on the QCM surface.[38] Therefore, QCM biosensors monitor the mass change of the immobilised recognition layer which occurs upon target hybridisation. These are label free biosensors.

Another mass-sensitive method has recently emerged which relies on the use of microcantilever sensors. Hansen *et al.*[39] have employed this method for the detection of a single-nucleotide polymorphism, by means of a gold-coated silicon AFM cantilevers onto which a thiolated 20-mer probe ssDNA was immobilised. Subsequently, hybridisation of the probe to the target was detected by the deflection of a laser beam reflected from the cantilever. These systems offer many advantages compared to conventional biosensors including reduction in the size of the device and high precision. However, the costs associated with such specialised instrumentation currently limit these methods to research laboratories.

In summary, numerous methods for the detection of nucleic acids have emerged and the limits of detection reached are today very low. The commercially available technologies

in use at the moment permit both the rapid detection of DNA and reproducibility of results. However, each of the mentioned techniques has its specific advantages and disadvantages and the ideal method is yet to be described. For a more comprehensive coverage of the state-of-play of DNA biosensors, the reader is referred to the review article by Sassolas *et al.*.[13]

1.3 PNA – PEPTIDE NUCLEIC ACID

1.3.1 Background

Nielsen *et al.* originally designed and developed peptide nucleic acid (PNA) as a reagent to sequence specifically target dsDNA as a mimic of triplex forming oligonucleotides. They believed that such synthetic oligonucleotides would bind as a third strand in the major groove of a DNA double helix *via* T·A–T and C^+·G–C Hoogsteen base pairing (*Figure 17*).[40-44]

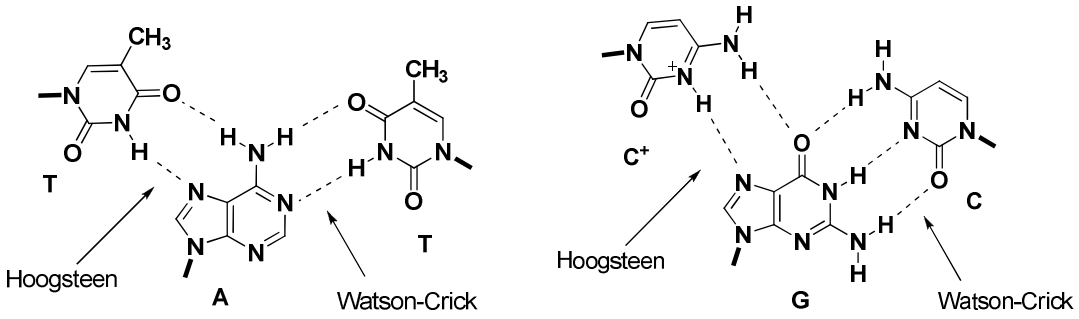


Figure 17 - T·A-T and C^+·G-C triplets involving Hoogsteen / Watson-Crick base pairs.

However, PNAs turned out to be a good structural mimic of DNA, attracting therefore wide attention in medicinal chemistry for development of gene therapeutic (i.e. antisense and antigene) and in genetic diagnostics. Here the entire negatively-charged, sugar-phosphate backbone of DNA was replaced with a neutral, "peptide-like", backbone consisting of repeating *N*-(2-aminoethyl)glycine monomeric units (*Figure 18*).[42, 43]

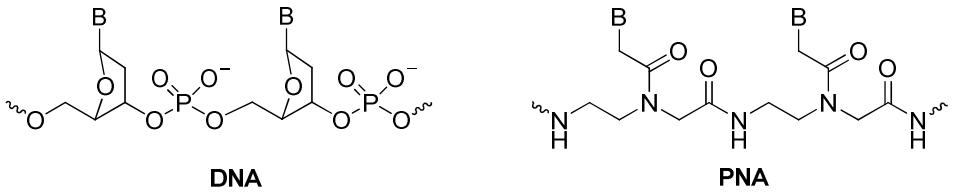


Figure 18 – The structure of PNA compared to DNA (B = nucleobase)

The four natural nucleobases were retained and attached to the central amine *via* an acetyl linker. This arrangement of atoms matches the '6 + 3' number of bonds arrangement found for DNA (i.e. in DNA six bonds separate each nucleobase unit and the distance between the backbone and the nucleobase is three bonds).[42, 43]

1.3.2 PNA hybridisation

PNA oligomers bind strongly and with higher sequence discrimination to complementary oligomers of DNA, RNA or another PNA, than their natural counterpart.[45] The resulting hybrid stabilities [quantified through the measurement of its melting temperature (T_m)] fall into the following order: PNA-PNA > PNA-RNA > PNA-ssDNA (> RNA-ssDNA > ssDNA-ssDNA). PNA-ssDNA hybrids are more stable than the corresponding DNA-DNA complexes, resulting in an increase in the melting temperature of approximately 1 ^{o}C/base.[45] PNA-RNA complexes are even more stable, showing an increase of about 1.5 ^{o}C/base[45] compared to their natural counterparts. PNA hybrids can be formed both in the antiparallel and the parallel configuration. The antiparallel complexes have the higher thermal stability, *ca.* 1-2 ^{o}C more stable, than the parallel ones.[45] The stabilities of PNA hybrids are, in contrast to hybrids between two anionic oligomers like DNA-DNA and DNA-RNA, relatively independent of ionic strength, due to the neutral character of the backbone. PNA also show greater sequence discrimination in recognising DNA. A decrease in T_m of between 8 and 20 ^{o}C for single base mismatches has been observed for a 15-mer PNA.[45]

Homopyrimidine and oligomers with high pyrimidine:purine ratio have been found to bind to complementary DNA sequence with a 2:1 stoichiometry *via* the formation of $(PNA)_2$/DNA triplexes using both Watson-Crick and Hoogsteen base pairing (*Figure 17* and *Figure 20)*. These triple helices are more stable than their corresponding DNA duplexes: the T_m value of the hybrid of the PNA H-$(T)_{10}$ with its Watson-Crick complementary deoxyribose $(dA)_{10}$ has been found to be 73 ^{o}C whereas the T_m value of the corresponding $(dT)_{10}$:$(dA)_{10}$ duplex has been measured to be less than 23 ^{o}C.[40]

Homopyrimidine PNAs can (but not always) bind to dsDNA with displacement of the Watson-Crick non-complementary strand to form a $(PNA)_2$-DNA triplex and a displaced strand analogous to a P-loop (*Figure 20*).[40, 46, 47] Homopurine PNA oligomers are capable of invading dsDNA but fail to form triplexes and result in less stable com-

plexes.[46] With a few exceptions,[47] strand displacement using mixed sequence PNAs has not been observed. It appears that they do not provide enough free energy upon hybridisation for this mode of binding. One possible solution to this problem is the simultaneous targeting of both dsDNA strands. This would require two complementary sequences of PNA which would "quench" one each other by forming very stable PNA-PNA duplexes.

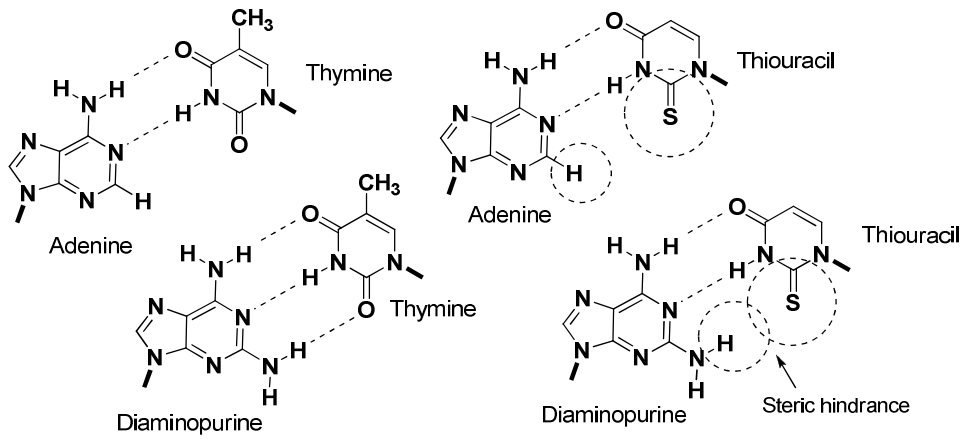


Figure 19 – Structures and interactions between pcPNAs with adenine and thymine.

However, Lohse *et al.*[48] have reported a solution of this limitation using pseudo-complementary PNAs (pcPNAs - here adenine and thymine have been substituted for diaminopurine and thiouracil, respectively - *Figure 19*). These recognise their DNA targets but not each other and therefore target dsDNA. This method proved successful and afforded double duplex strand invasion complexes whereby each of the strands in the DNA is bound to pcPNA.

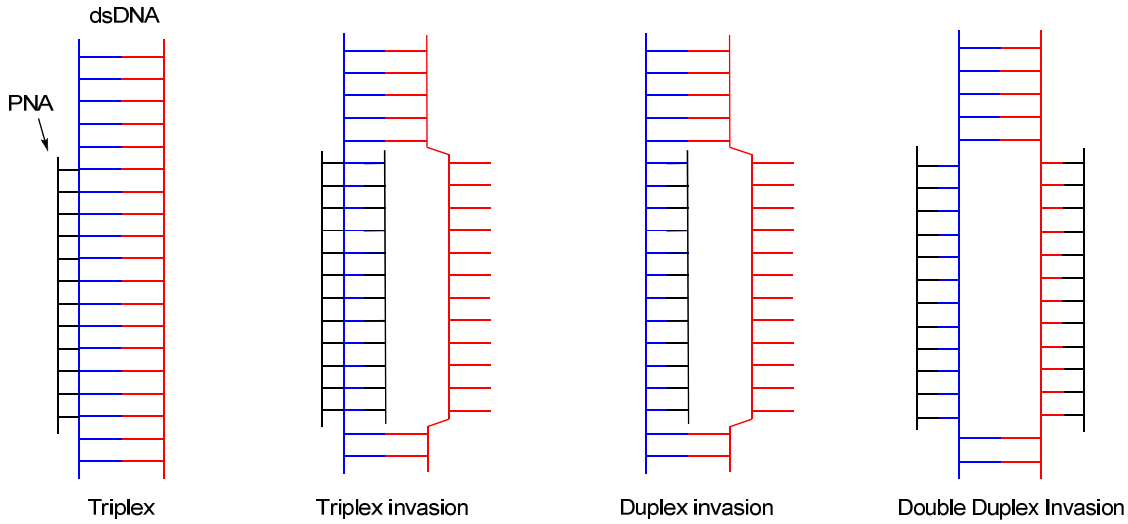


Figure 20 - Modes by which PNA can bind to dsDNA.

1.3.3 PNA in therapeutic strategies

Owing to their unique properties briefly listed in the previous section, PNAs represent attractive candidates for the next-generation of genetic therapeutic drugs designed to interfere selectively with gene expression. The use of PNA oligomers in antisense and antigene biotechnologies holds considerable promise.

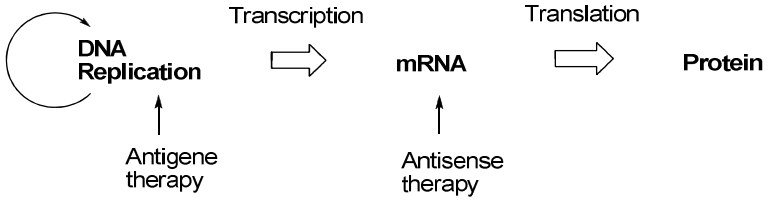


Figure 21

In principle, two strategies can be adopted (*Figure 21*): (1) PNA can be designed to recognize and hybridise to complementary sequences in a specific gene and therefore interfere with the transcription of this gene (i.e. antigene therapy); and (2) a PNA oligomer can be designed to target mRNA and thereby inhibit its translation into proteins (i.e. antisense therapy).[49, 50]

The potential of PNAs as both transcription and translation inhibitory agents has been shown through a number of *in vitro* experiments. Translation arrest requires a $(PNA)_2$-RNA triplex, thus a homopurine target of 10-15 bases. Mixed sequence duplex-forming PNAs have been shown to exhibit antisense activity only when targeted to the AUG (initiation) region, whereas similar PNAs had no effect when targeting the coding region.[51] Since PNA-RNA duplexes are not substrates for ribonuclease H (RNase H), the resulting antisense effect is based on the steric blocking of RNA processing or translation.[51] Mologni *et al.* have reported the effects of three different types of antisense on the *in vitro* expression of the PML/RARα gene.[52] The first one was complementary to the first AUG site. The second could bind to a sequence in the coding region that includes the second AUG, the starting site for the synthesis of an active protein. The third PNA was targeted against the 5′-untranslated region (UTR) of the mRNA, the point of assembly of the translation machinery. Limited effects were obtained when these PNAs were used separately, but they provided an efficient inhibition of translation when used together. Almost total inhibition was obtained at concentration between 0.6 and

0.8 μM.[52] These data show that an effective translation inhibition can be achieved by combining PNA targeting at 5′-UTR and AUG regions.

PNAs have also been shown to arrest transcriptional processes thanks to their ability to form a staple triplex structure and a strand-invaded or strand displacement complex with DNA.[53] An example has been reported by Lee *et al.*,[54] who have demonstrated that PNA complementary to the primary site of the HIV-I genome can completely block priming by $tRNA_3Lys$. Consequently, *in vitro* initiation of reverse transcription by HIV-1 RT was blocked. These results show that PNAs targeted to various critical regions of the viral genome can have strong therapeutic potential for interrupting the replication of HIV-I.[54]

Unfortunately due to the poor uptake of PNA by living cells, further investigations *in vivo* have been held back.[55] However, a variety of potential cellular delivery systems are currently being investigated in order to overcome this difficulty. These involved the modification of PNA through conjugation to lipophilic moieties and the conjugation to cell-membrane crossing peptides.[55] Regrettably, to date these have not shown the break-through expected.

1.3.4 PNA as a diagnostic tool

Thanks to their unique hybridisation properties, PNAs have found applications in the field of detection of point mutation responsible for genetic diseases (see Section 1.1.5). PNAs have been used in pre-diagnostic tools such as PCR clamping. This application relies on the fact that PNAs cannot act as primers in the polymerase chain reaction (PCR).[56, 57] PCR is widely used as amplification tools prior to performing nucleic acid detection, when the quantity of DNA in the sample is too small for analysis. PCR requires two oligonucleotide primers, one to hybridise each DNA strand. Their binding sites flank the sequence to be amplified. A PCR cycle consists of the DNA first being denatured at high temperature (90 ^{o}C) and then allowing the primers to anneal at a lower temperature (50 ^{o}C). Subsequently, the polymerase chain reaction, carried out by the enzyme DNA polymerase, takes place at an intermediate temperature (70 ^{o}C). This cycle is repeated up to 30 times in order to greatly increase quantities of the desired DNA sequence. Unfortunately, the general applicability of this approach is limited by the fact

that the majority of primer-template mismatches have no significant effect on the amplification process.[58]

In order to improve this situation, Ørum *et al.*[56, 57] have developed a method that enhances the specificity of the PCR reaction by targeting the initial step involved in non-specific amplification, known as PCR clamping (*Figure 22*).

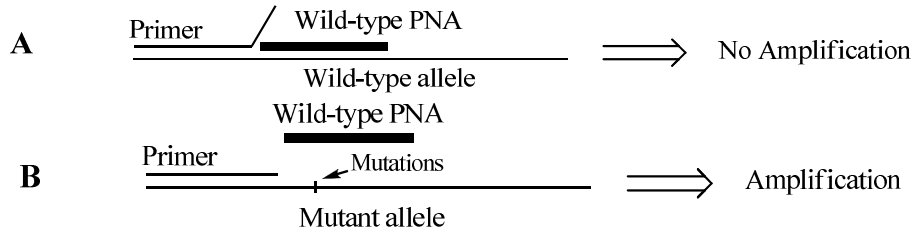


Figure 22 – Schematic illustration of PCR clamping method.

A PNA oligomer complementary to the wild-type genetic nucleic acid sequence under investigation is added (**A** – *Figure 22*). This hybridises to the target and sterically hinders annealing of the primer. This excludes the wild-type sequence from successful PCR amplification. However, for mutated sequences (**B**) the primer is able to outcompete the PNA and binds to the DNA. Therefore, allowing for preferential amplification of mutated sequences only. This method has been shown to be successful at both the level of single[56] and multiple[59] base mutations.

The overall performance of the DNA biosensors using ssDNA probes described in Section 1.2 has been found to be dependent upon experimental variables that influence the hybridisation efficiency between the target and the analyte (e.g. temperature, probe length, ionic strength).[60] Even under strictly controlled conditions, in the presence of single-mismatch ssDNA probes are still capable to hybridise to the analyte, therefore leading to a false result. They are therefore not very efficient for the detection of disease related point mutation. These limitations have been overcome by the introduction of PNA as a recognition element in nucleic acid biosensors and have enhanced the power of DNA biosensors. As mentioned earlier in this section, PNA have a higher affinity towards DNA than DNA does and the presence of a single mismatch is more destabilizing in PNA/DNA duplexes than it is in DNA/DNA duplexes. These properties have allowed the development of more sensitive methods for the detection of mutations and single nucleotide polymorphisms (Section 1.1.5).

Like it is the case with ssDNA probes, the immobilisation step of PNA probes is very important and must result in readily accessible PNA probes surface layer. The same kind of immobilisation methods are used for PNA probes as it was the case for ssDNA probes (Section 1.2). Similar transducers (i.e. electrochemical, optical and mass-sensitive) have also been used for monitoring the hybridisation between PNA probes and nucleic acids targets. The introduction of PNAs as probe has also opened the way to simpler indicator-free detection schemes. For example, electrochemical strategy could be used for detecting the sugar moiety of the hybridised DNA target (since PNA do not have a sugar backbone). Also PNAs do not have phosphorus, so the phosphorus signal can be used for detecting the hybridised DNA. Laser-based fluorescence detection, involving two different PNA probes labelled with a different dye has been used to screen mixtures of oligonucleotide sequences.[61]

Wang *et al.* have extensively studied the detection of the frequent point mutation found on p53 gene.[7] These researchers have used two 17-mer probes with the same sequence, one ssDNA and one PNA. They also used three DNA targets, Target I is a 17-mer and is complementary to the probes while Target II (also a 17-mer) has a one-base mismatch, finally Target III (a 30 mer) is an extension of the 17-mer DNA at both ends (*Figure 23*).

PNA probe: H-GGG GCA GTG CCT CAC AA-LysNH_2
DNA probe: 5′-GGG GCA GTG CCT CAC AA-3′
Target I: 5′-TT GTC AGG CAC TGC CCC-3′
Target II: 5′-TT GTG AGG C<u>G</u>C TGC CCC-3′
Target III: 5′-CGG AGG <u>TTG TGA GGC ACT GCC CC</u>C ACC ATG-3′

Figure 23

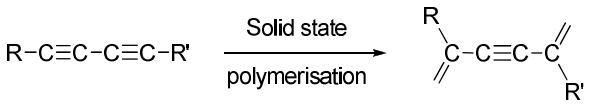
The hybridisation of probes to targets was monitored using the sensitive constant-current chronopotentiometry (i.e. electrochemical) with a $Co(phen)_3^{3+}$ redox indicator on a carbon paste electrode. The results obtained showed that the PNA probe was, as expected, more discriminating towards the single mismatch present on Target II than the corresponding DNA probe.

This remarkable mismatch discriminating capability of PNA based biosensors has also been illustrated in connection with the use of quartz crystal microbalance transducer.[62] Finally, PNA is easily synthesised on automated instruments and this offers the possibility to prepare high density micro-arrays for high throughput experiments.[63] In the overall, PNA has more applicability as probes than their natural counterpart.

1.4 POLYDIACETYLENE (PDA) MATERIALS

1.4.1 General background and polymerisation

It has been known for over a century that on application of heat, γ or UV-irradiation, diacetylenes undergo a colour change in the solid state.[64-67] However, the chemical change was not understood until 1969 when Wegner *et al.*[64] suggested a 1,4-addition of successive diacetylene molecules affording a polymer with a fully conjugated backbone.[66]



Scheme 1

The polymerisation can be initiated by photochemical, thermal or high-energy irradiation of the monomer crystals. Polymerisation proceeds within monomer crystals and is completely controlled by the packing of the monomer and is therefore termed topochemical.

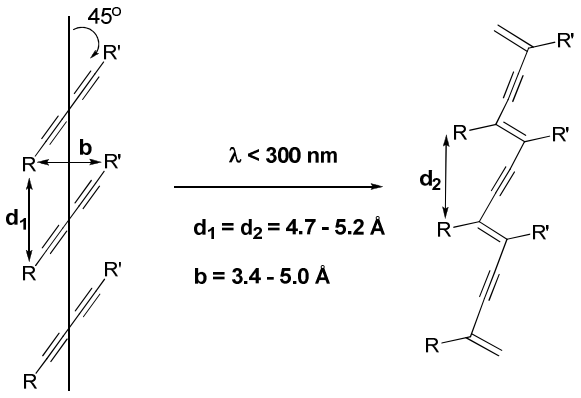


Figure 24

For polymerisation to occur, the diacetylene monomers must be stacked so that one triple bond approaches the adjacent triple bond system at a distance lower than 4 Å.[67] Po-

lymerisation proceeds by successive titling of each molecule along the ladder with a minimum of movement of all atoms from their lattice site, this is known as the "least motion principle". The angle between the diacetylene rod and the translational axis has to be of *ca.* 45° so that the specific volume and lattice symmetry can be retained.[68] It has been observed that the distance obtained between the two side chains (d_1 and d_2) is very similar before and after polymerisation and this for a variety of studied polydiacetylenes (*Figure 24*). The polymer obtained from this reaction is usually highly coloured, with shades from blue to red to yellow.

When PDA monomers have amphiphilic character, the monomers align and self-assemble in aqueous solution. Amphiphilic diacetylenes are made of two parts: a polar headgroup and a hydrophobic lipid tail containing the diacetylene moiety. The tail itself is separated in three domains: the diacetylene group, the spacer between the headgroup and the diacetylene, and the terminal alkyl chain or lipid-chain (*Figure 25*).[69]

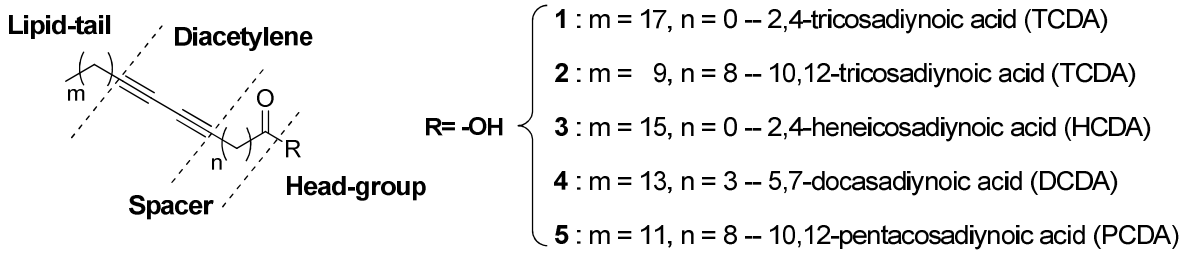


Figure 25 – Diagram of a typical amphiphilic diacetylene and examples.

Each element sited above plays a role in determining the shape that the self-assembled material takes. For example, the headgroup chirality strongly influences colloidal structure; amphiphilic monomers with achiral headgroups usually form spherical liposomes, whereas chiral amphiphiles form structures such as helices and tubules. The number and length of the tails largely controls the range of conditions under which an amphiphile will undergo self-assembly.[69] Tieke *et al.*[70] showed that the relative position of the diacetylene plays a significant role in the polymerisation and on the colour of the PDA obtained. When the diacetylene was moved closer to the head-group the polymerisation was reduced and the generated polymer was red or yellow rather than blue. For example, 8,10- (m = 13; n = 6 – *Figure 25*) and 10,12-PCDA (**5**) give a blue polymer under UV irradiation while 6,8-PCDA (m = 15; n = 4) gives a red polymer and 4,6-PCDA (m = 17; n = 2) gives a weakly yellow one.[69] It was also reported that a shorter spacer

leads to better packing of the amphiphiles but poorer polymerisation, and that an even spacer was found to give PDA with more blue character.[71]

A large variety of forms and shapes have been prepared, depending on the application they were required for. These structures have been divided into five categories:[69] Langmuir, Langmuir-Blodgett (LB), Langmuir-Schaefer (LS) films; self-assemble monolayers (SAMs); multilayer coatings; colloids; and immobilised colloids (*Figure 26*).

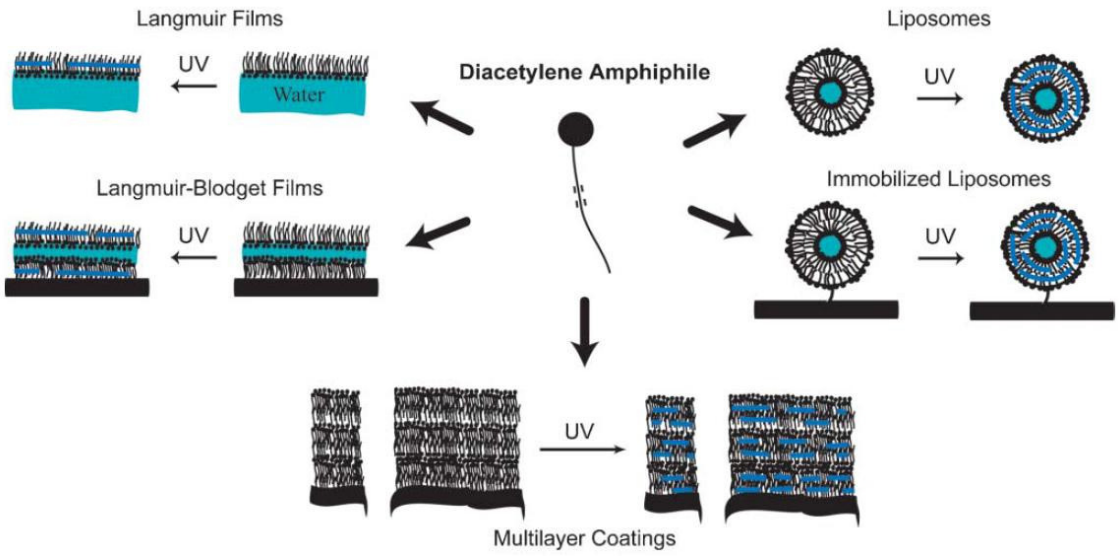


Figure 26 – Variety of forms obtained for diacetylene amphiphiles.
Reproduced from ref. 69 with permission of The Royal Society of Chemistry

Langmuir films are formed by spreading amphiphilic molecules at an air/water interface and compressing them into close-packed arrays. This form of PDA assemblies have been used for a variety of detection platforms.[72] These films can then be studied directly at the air/water interface or transferred on solid support to make LB or LS films. Langmuir films are useful forms for spectroscopic and microscopic studies but their preparation requires skilled workers and sophisticated instruments. Jelinek *et al.*[73] prepared Langmuir films by mixing diacetylene 10,12-TCDA (**2** – *Figure 25*) and a phospholipid. These researchers noticed that the two components phase separated and that this segregation could be used for the incorporation of cell membrane components (i.e. proteins). This type of assembly has been used to study the biomimetic character of PDA-based materials.[74]

PDA liposomes (which can be called vesicles) have also proved to be a popular platform for bioassays development owing to their biomimetic nature.[68] The words lipo-

somes and vesicles have been interchangeably used and broadly referred to any structure composed of amphiphilic molecular bilayers that enclose a volume. Liposomes are formed by hydration of the appropriate amphiphile material in water or buffer employing a variety of methods such as sonication, solvent injection, or simple shaking.[75] The diacetylene-containing lipids must hydrate and organize themselves into the correct packing and orientation in order to undergo the subsequent 1,4-addition to the conjugated polymer. Thanks to their shape, liposomes have been extensively used as biomimic of cells for membrane interactions studies. It is possible to incorporate up to 70% of non-polymerisable lipids in diacetylene liposomes to give them the biomimetic character, with a membrane fluidity comparable to cell membranes, without interfering with the polymerisation of the diacetylene components of the PDA.[76-78] Liposome formation can be directly verified by transmission electron microscopy (TEM) or indirectly by the optical detection of a resulting blue colour; the polymerisation reaction itself is the evidence that an ordered lipid assembly has formed (but the converse is not always true). The formation of a coloured polymer can, therefore, be seen as a quick (and cheap) "benchtop" test of vesicle formation.[68]

Thus, it is not surprising that this class of PDA-assemblies, coupled to their chromic properties, has led to their investigations for use in a wide range of biosensing applications.[79] The most promising of these applications will be discussed later in more details in Section 1.5.

1.4.2 Optical properties of PDAs

As mentioned earlier the polydiacetylene polymer is formed by 1,4-addition of diacetylenic monomers, initiated by UV irradiation and the resulting polymer is intensely coloured, and typically deep blue. This deep colouration arises from π-to-π^* absorption within the linear π-conjugated backbone (due to alternating triple-bond/double-bond backbone structure). Absorption spectra of PDA materials typically show several peaks with maxima at *ca.* 650, 550 and 500 nm. These peaks are associated with the blue (650 nm) and the red forms (or phases) of PDA (550 and 500 nm).[80] The appearance and the ratio of absorbance peaks for given PDA material depend on the structure of the polymer side chains, the extent of the polymerisation and the environment of the polymer. These polymers have been found to undergo colour transitions (usually from blue to red) in response to environmental stimuli: organic solvents (solvatochromism)[81, 82],

heating (thermochromism),[83-87] mechanical stress (mechanochromism)[72, 88] and ligand-receptor interactions (affinochromism or biochromism).[76, 89-94] This colour transition is believed to be caused by structural perturbations that give rise to a gauche-*trans* conformational transition of the PDA backbone, resulting in shortening of the conjugation network and absorption at a shorter wavelength.[95] The mechanism of chromic change has been studied by many research groups and it is now recognized that, upon polymerisation, the alkyne carbons of the diacetylenic monomer change hybridisation from sp to sp^2. The preferred bond angle also changes from 180^o to 120^o. This constriction results in an accumulation of strain in the material as polymerisation proceeds. The packing of the side chains creates a barrier that prevents the backbone from adopting a more relaxed form. Chromic change is, therefore, seen as a way to overcome the energy barrier to reorganization of the polymer backbone created by the side chains packing. An increase of the temperature augments the motion of the side chains and allows changes in their packing. pH changes alter the hydrogen-bonding in head-groups allowing the methylenes between these head-groups and the backbone to reorganise. In the case of target binding, a disruptive effect may induce changes in the head-group's hydrogen-bonding and alkyl chain packing. These conformational changes, thus, allow the backbone to adopt a more relaxed conformation that leads to the colour change.[69]

The extent of the 'blue-to-red' transition exhibited by PDA materials upon exposure to external stimuli can be quantified *via* the colourimetric response (% CR). This term was first described by Charych et *al.*[90] and is defined in *Equation 2*:

$$CR = \frac{B_v - B_0}{B_0} \times 100\ \% \qquad \textit{Equation 2}$$

B is the percentage of blue component in the material calculated from A_{blue}, the absorbance of the "blue" component (*ca.* 650 nm) and A_{red}, the absorbance of the "red" component (*ca.* 550 nm) (*Equation 3*):

$$B = \frac{A_{blue}}{A_{blue} + A_{red}} \qquad \textit{Equation 3}$$

B_0 is the percentage of "blue" of the PDA material in the absence of the external stimuli while B_v is the percentage of "blue" after exposure to the external stimuli. [Note: "blue" and "red" refer to the visual appearance of the material, not its relative absorbance.]

Most examples of biosensors based on PDA structures that have been reported in the literature show an irreversible blue-to-red colour change. That is, on removal of the external stimulus, the original colour of the PDA material does not return. However, reversibility is a mandatory feature for a reusable biosensor. Jonas and Charych[96] have prepared hydrazide-modified diacetylene lipids **6** and **7** (*Figure 27*) and constructed PDA liposomes. These assemblies were reported to exhibit a reversible colour change when the pH of the solution was cycled between acidic and basic conditions.

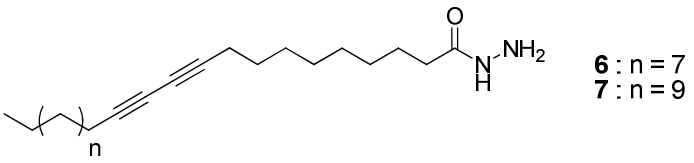


Figure 27

More recently Ahn and Kim[84] have prepared the PDA assemblies from diacetylene **8** (*Figure 28*) that showed reversible colour changes as the temperature of the solution was cycled. They demonstrated that the double hydrogen-bonding system amongst the head-groups provides the film with the ability to recover its initial molecular organization once the stimuli are removed.

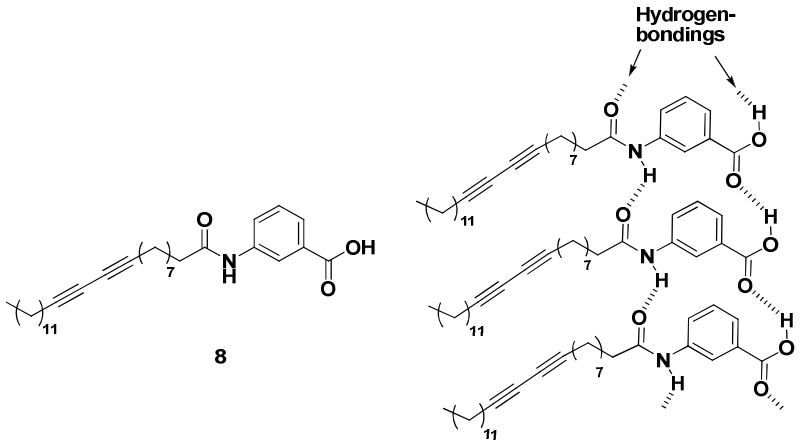


Figure 28

Thus, it was reasoned that the irreversibility commonly noticed in PDA assemblies, was due to the hydrogen-bonding involved within the head-groups that are not strong enough to maintain and recover the original molecular organisation that the structure had before the application of the stimulus. However, no reversibility in the colour change has been demonstrated on the removal of a biological analyte.

In addition to the remarkable and readily detectable chromic properties, PDA materials in their "red" phase have been found to be fluorescent whilst "blue" PDA assemblies are not. "Red" PDA liposomes or films can be excited with wavelengths above 450 nm and this leads to the emission of two broad fluorescent peaks at *ca.* 560 nm and 640 nm.[69] This feature is highly attractive as such a system would generate a fluorescence (i.e. turn-on type) signal which is easier to detect rather than experiencing a loss or decrease of fluorescence (i.e. turn-off type).[80] The mechanism giving rise to this fluorescence property is believed to be due to energy shifts in the lowest-lying excited states; the lowest excited state in the "blue" PDA has the same symmetry as the ground state (A_g) which is a dipole-forbidden transition. In "red" PDA, the lowest excited state has B_u symmetry, allowing radiative decay.[80] It was shown that the fluorescence signals obtained from "red" PDA materials are detectable using a commercially available fluorescence microscope. To date very few research groups have reported the use of this property for the development of PDA-based biosensors.[97]

The optical properties of PDA materials and their environmental susceptibility have been the basis for the majority of PDA-based detection systems. PDA films, liposomes and immobilised liposomes have proven to be effective biosensors for the detection of various biological entities as schematically represented in *Figure 29*.

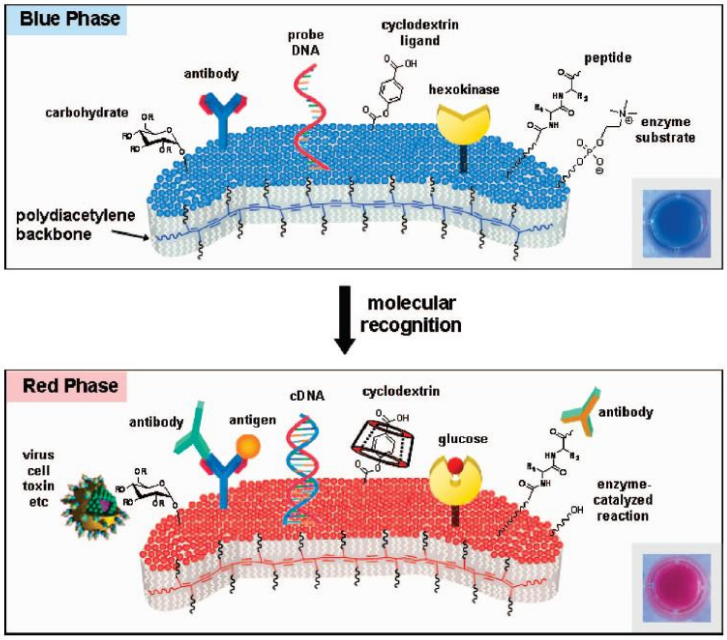


Figure 29 – Schematic representation of surface ligands interactions developed in colourimetric PDA sensors.Reproduced with permission from ref. [84]. Copyright 2009 American Chemical Society

1.5 BIO-DETECTION APPLICATIONS OF PDAS SENSORS

PDA-based biosensors have been intensively investigated for the detection of various biological molecules including viruses,[90, 91, 98] proteins,[76, 99] peptides,[100] oligosaccharides,[101] bacteria,[102, 103] lipophilic enzymes,[89, 104] ions,[94] and drugs.[105, 106] A selection of the most promising PDA biosensors studied to date is covered in more detail in this section.

1.5.1 The use of PDA-based biosensor for the detection of viruses

The detection of viruses, such as the influenza virus, requires a bioassay both sensible and rapid in order to give the correct treatment to the patient. Influenza is an enveloped RNA virus, with two glycoproteins at its surface, hemagglutinin (HA) and neuraminidase (NA). These two proteins play important roles in infecting host cells. Influenza viruses are classified in three categories (A, B, C) following the type of nucleoprotein present in the virus. Influenza A is further divided into subtypes based on serological cross-reactivity of HA and NA antibodies and they are expressed as a combination of 15 types of HA (H1-H15) and nine types of NA (N1-N9), for example H3N2 or H1N1.[107] This disease is responsible for hundreds of thousands of worldwide deaths each year and is circulated very rapidly through the population. Although, the methods currently available for the detection are very precise, they are also lengthy and require labour-intensive laboratory procedures as well as trained personal to carry them out. The period of incubation for the influenza viruses is 24 hours. If the patient misses the 48 hour window, he will not be able to take advantage of the potent antiviral drugs (i.e. Zanamivir and Oseltamivir) currently available on the market. Therefore, a clinical diagnosis tool providing on-site result is highly demanded, for fast diagnosis of the patient.[108]

In 1993, Charych *et al.*[90] described the first biosensor based on a polymerized bilayer assembly capable of detecting influenza A virus. In this work, their research established the whole field of PDA-based colourimetric biosensors. The bilayer developed was composed of a self-assembled monolayer of octadecyltrichlorosilane (OTS) and a LB monolayer of functionalised PDA. This film was designed to undergo a similar colourimetric response to that which occurs through solvatochromism and thermochromism. The bilayer incorporated both a molecular recognition binding site and a detection element, to enable rapid and quantitative detection by visible absorption spectroscopy.

The recognition site introduced in the assembly was in the form of a sialoside lipid monomer (**9** – *Figure 30*). This monomer contains a carbon-linked sialic acid head-group which provides the molecular recognition site for the virus. In nature, the virus binds to terminal α-glycosides of sialic acid of cell surface glycoproteins and glycolipids *via* hemagglutinin, thus initiating cell infection by the virus. In this work, Charych *et al.* replaced the naturally occurring oxygen glycoside with a carbon-linked in order to avoid hydrolysis by neuraminidase, an enzyme present on the surface of the virus. Earlier work showed that this replacement did not affect the binding affinity of hemagglutinin.[109]

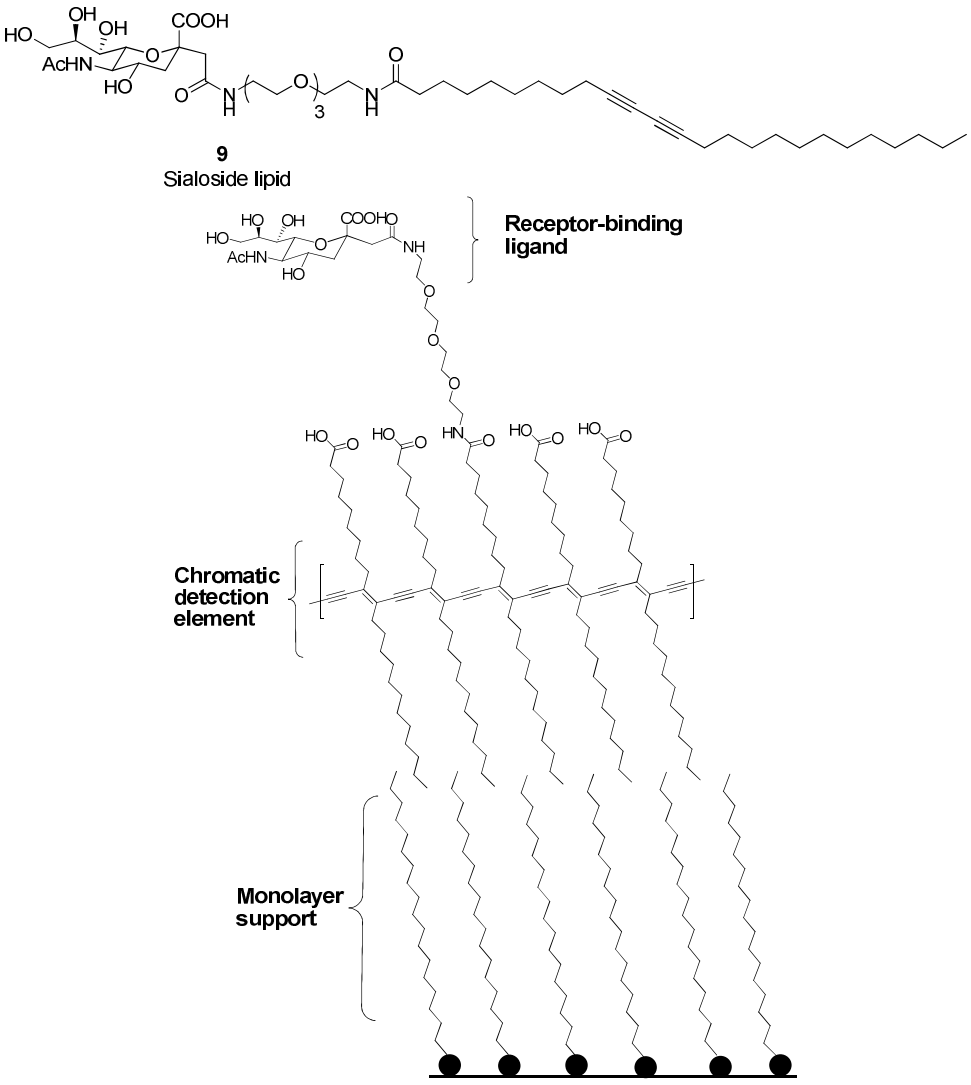


Figure 30 – Schematic diagram of the polymerised bilayer assembly.

Therefore, a solution made from mixtures of 2 to 5% of sialoside lipid **9** and matrix lipid 10,12-PCDA (**5**) was spread on the water surface of a standard LB trough. The sialoside lipid was uniformly dispersed within the matrix lipid which allows optimum

binding of the virus.[90] The resulting mixed monolayer was then compressed and polymerized by UV irradiation. The floating polymer assembly was lifted by the horizontal touch method onto a glass slide previously coated with self-assembled monolayer of OTS.[110] The resulting bilayer assembly presented an array of carbohydrate ligands at the surface, linked to the polydiacetylene reporter unit. The tetraethylene glycol spacer served to extend the carbohydrate molecular recognition site beyond the carboxylic acid headgroups of the matrix lipid (*Figure 30*).

The bilayer appeared as a blue film with an absorption maximum of 620 nm and a weaker absorption at 550 nm (solid line *Figure 31*). When the film was incubated with X31 influenza A virus, the binding of the viral hemagglutinin to the sialic acid residues on the surface resulted in a blue-to-red colour transition. The maximum at 550 nm increased with a decrease in the maximum at 620 nm (dotted line *Figure 31*). No colour change was observed when the blue film was incubated with a blank solution of PBS buffer. This result demonstrated that polydiacetylene colour change arose from affinity binding (affinychromism).[90] From the absorption spectra (*Figure 31*), Charych *et al.* were able to quantify the degree of colour change using the principle of colourimetric response (CR), explained earlier in Section 1.4.2.

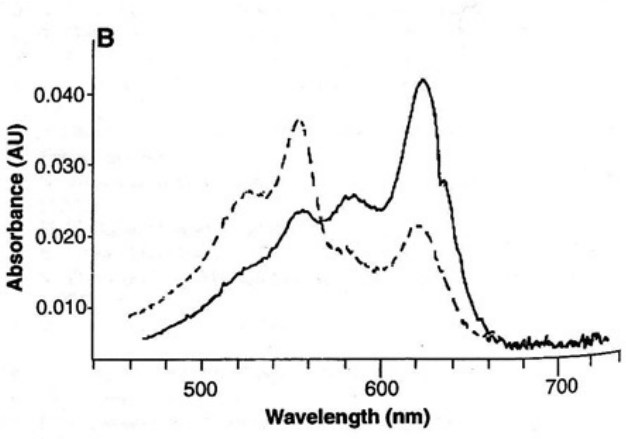


Figure 31 – Visible absorption spectrum of bilayer assembly before (solid line) and after (dashed line) viral incubation - From Ref. 90

They found that the CR was directly proportional to the quantity of influenza virus employed[90] and that the colour change was specific to interactions with the sialic acid residue. In order to prove the latter, a number of competitive inhibition assays were carried out. For example, the bilayer assembly and the virus were incubated together in the presence of a known binding inhibitor, α-methyl-*N*-acetyl neuraminic acid (**10** – *Figure*

32). In this case, no CR was observed. In order to confirm that this inhibition had been specific, the film and the virus were incubated together in the presence of two known non-inhibitor compounds, β-methyl-*N*-acetyl neuraminic acid (**11**) and α-glucose (**12**). This time, for both cases, a similar colour change was observed as when the PDA assembly was incubated with the virus alone.

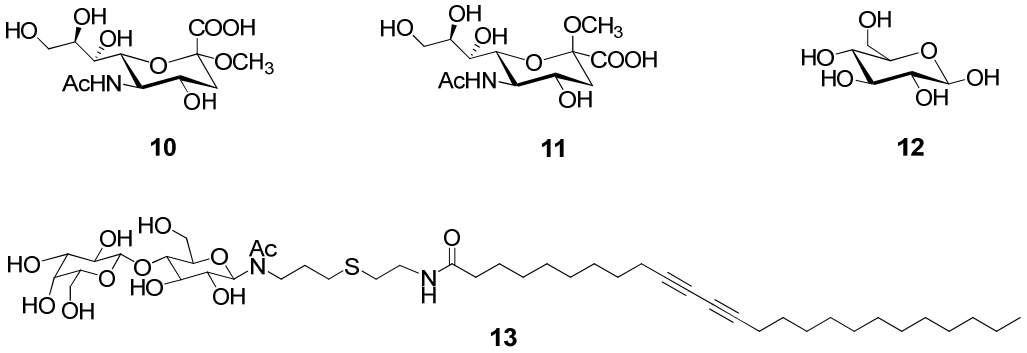


Figure 32 – Compounds used for competitive inhibition experiments.

The last experiment which was performed to show that specific adhesion of the virus to the PDA film had provided the colour change involved the use of a bilayer assembly in which the sialoside lipid (**9**) had been replaced by a lactose lipid (**13** – *Figure 32*). Lactose does not bind to hemagglutinin lectin. In this case, only a small CR (2 to 4%)[90] was obtained after incubation of the film with the influenza virus. Finally, incubation of the original sialic acid derived assembly to a concentrated solution of bovine serum albumin failed to show any significant CR. These last two studies, when taken together, clearly indicated that non-specific adsorption of virus or protein to the film was not responsible for the dramatic CR observed in specific upon exposure of the sialic acid-containing PDA film to influenza A. From these investigations, Charych and co-workers concluded that the observed "blue-to-red" transition had to have been triggered by the specific binding of the virus to the sialic acid "head-group" leading to changes in the PDA backbone conformation.

Following this finding, the same research group investigated the possibility of constructing a similar influenza virus biosensor where this time the assembly would take the form of liposomes (*Figure 33*).[91] As mentioned earlier (Section 1.4.1), liposomes form when phospholipid or amphiphilic compounds are suspended in water and are sonicated (i.e. agitated by ultrasonic vibration). Thus, Charych *et al.*[91] prepared liposomes using the same mixture of sialic acid-functionalised diacetylene lipid **9** (5-10%)

and matrix lipid **5** as in their previous work on PDA thin films. To do that, the two lipids were first mixed in chloroform in a test tube followed by evaporation of the solvent with a gentle stream of nitrogen; this left a thin film of lipid on the glass surface. To this was then added an appropriate amount of de-ionised water, so that the total lipid concentration was 1 mM. The mixture obtained was then heated for 15 minutes at 80 ^{o}C whilst being subjected to sonication. The warm solution was subsequently filtered through a 0.8 μm nylon filter to remove undispersed lipid and finally cooled overnight at 4 ^{o}C. Subsequently, this liposome solution was purged with nitrogen gas for 5 minutes and allowed to warm to room temperature. Finally, polymerisation was achieved by irradiation using a UV lamp (254 nm) positioned at a distance of 3 cm, for varying periods of time. Irradiation for 5 to 10 minutes was found to result in blue coloured liposomes whereas irradiation for longer (up to 30 minutes) resulted in purple solutions.[91]

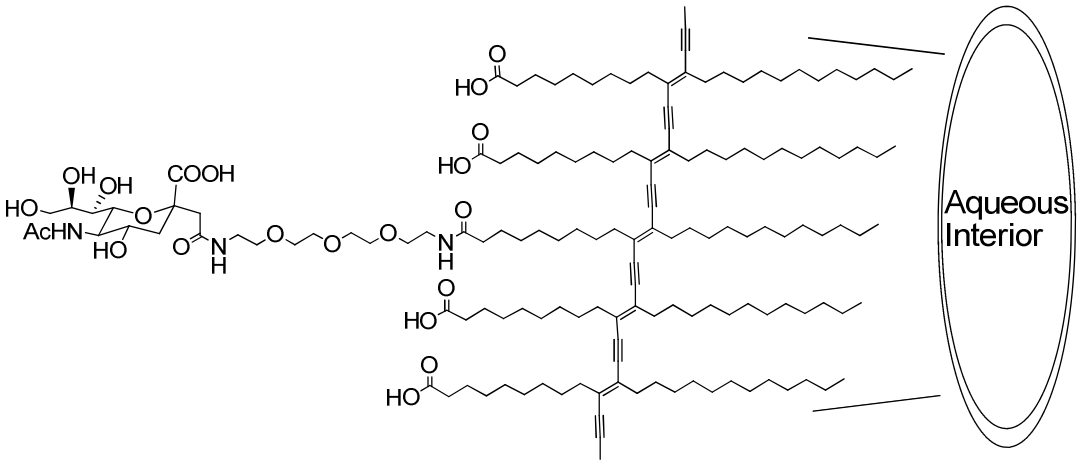


Figure 33 – Schematic diagram of the polymerised liposome assembly.

These aqueous solutions of sialic acid functionalised PDA liposomes were then examined for their abilities to detect the influenza A virus. It was found that when the influenza A virus was added to the liposomes in PBS buffer, the solution immediately changed to a pink or orange colour, depending on whether the initial liposome preparation was blue or purple, respectively. The CR was determined for the different preparations and it was shown to be dependent on the amount of UV irradiation used for the polymerisation. The CR was 47% for the blue liposomes whereas the purple liposomes gave rise to a CR of 87%. It was, therefore, postulated that the enhanced sensitivity of the purple PDA liposomes was due to increased polymer content, as suggested by their higher optical density. Identical inhibition control experiments were carried out as has been described previously for the analogous PDA films and again the results indicated

that the observed CR was triggered by specific binding of the virus to the sialic acid head-groups.

In addition to the work reported above on the use of PDA assemblies for the detection of influenza A, Jiang and co-workers[111] have recently shown that PDA-based liposomes can be effectively used for the identification of avian influenza (H5N1) hemagglutinin HA1. In this case, the recognition element prepared consisted of active sialic acid-β-glucoside (G1) and inactive lactose-β-glucoside (G2) receptors (either used separately or in conjugation). Both were attached *via* a tetraethylene glycol spacer to a C_{12} hydrocarbon chain. These lipids were then physically embedded at different ratio (G1 only, G1+G2 and G2) within a matrix lipid made from 10,12-PDCA (**5**) and DMPC (dimyristoylphosphatidylcholine), through non-covalent interactions. They found that larger concentration of DMPC yielded more sensitive liposomes, up to a certain threshold value. Jiang *et al.* observed that G1-liposome solutions underwent a blue to purple transition as the HA1 concentration reached 20 ng/mL, whereas the colour of the G2 analogue remained blue, meaning that no conjugation occurred. Finally, liposome-(G1+G2), especially at 5:5 ratios, caused a remarkable colour change from blue to red even at low concentration of HA1 (*ca.* 10 ng/mL), showing a strong synergistic effect for the recognition of virus HA1. Although these tests have yet not been carried out *in vivo*, these findings demonstrated the possibilities of using PDA liposomes for future detection of pandemic viruses.

1.5.2 The use of PDA-based biosensor for the detection of toxins

A couple of years after the development of the influenza A virus biosensor, the group of Charych published results on the preparation and use of PDA assemblies (films[98] and liposomes[103]) incorporating receptor molecules for the detection of toxins. Gangliosides are part of a family of molecules that reside within the cell membranes of neurons. They are lipid molecules which bear a carbohydrate group covalently attached to the extracellular end of the lipid unit. The lipid chain anchors the carbohydrate head-group into the cell membrane. Two representative members of this family are the G_{M1} (**14**) and G_{T1b} (**15**) gangliosides (*Figure 34*). The G_{M1} gangliosides, present on the surface of intestinal cells, are the primary target of cholera toxin, the neurotoxin responsible for the disease cholera. The G_{T1b} gangliosides are located at the neuromuscular junction and are the primary target of botulinum neurotoxin, the neurotoxin responsible for botulism.

Since here natural gangliosides already contained a lipid functionality, Charych *et al.*[98, 103] demonstrated that they could be directly incorporated into artificial PDA assemblies through non-covalent interactions in a manner analogous to the heterogeneous mixing of molecules in cell membranes.

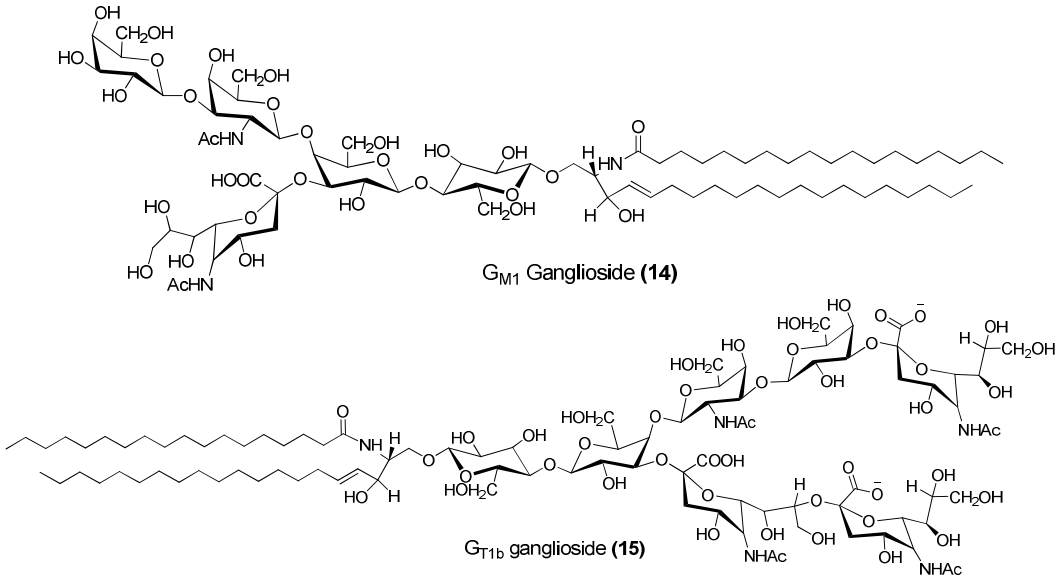


Figure 34 – Structures of G_{M1} and G_{T1b} ganglioside used by Charych et al.

The chromatic unit of the neurotoxin sensor was in his core composed of PDA and of either sialic acid PDA (**9**) or lactose-derivatized PDA (**13**) which acted as promoters (*Figure 35*).[98] The presence of the "promoter" was found to be essential for the production of a colourimetric response. It was postulated that the "promoter" PDA lowers the activation barrier for the chromatic transition, by changing the lipid packing and altering the effective conjugated length of the PDA backbone. It may also provide a connection between the non-conjugated receptor and the conjugated backbone, thereby enabling binding of the neurotoxin to the ganglioside headgroup and induce the associated colourimetric transition. The LB films contained 5% of the ganglioside lipid G_{M1} (**14**) or G_{T1b} (**15**) and either 5% of sialic acid-functionalised lipid (**9**) or 2% of lactose-functionalised lipid (**13**). If higher concentration of gangliosides were used, then the polymerisation was hindered by steric effects and films of poor quality were formed.[98]

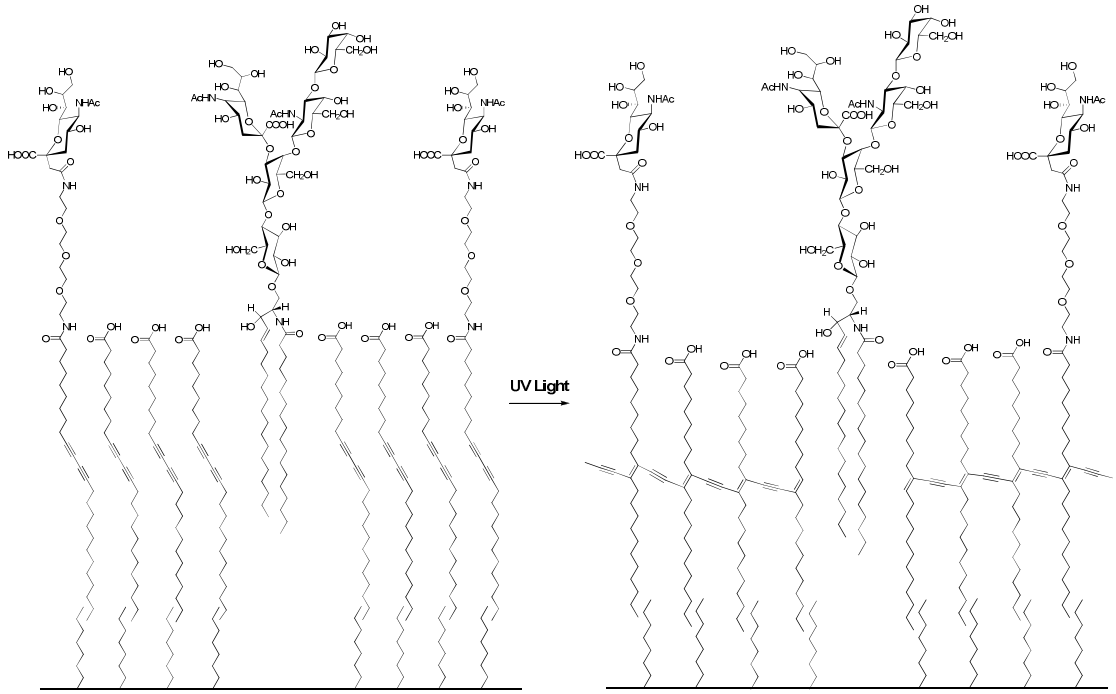


Figure 35 – Schematic representation of LB film assembly containing G_{M1} and sialic acid-functionalised diacetylene lipid.

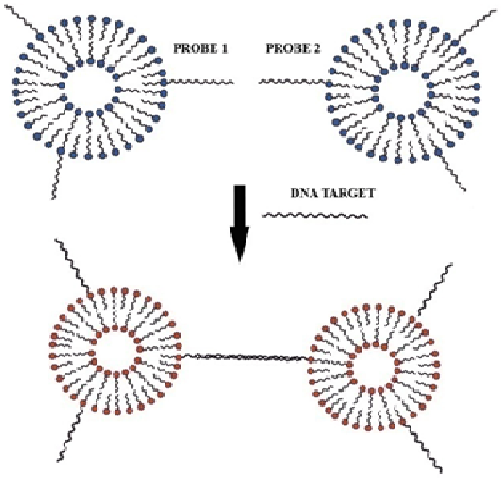
By analogy with the influenza virus biosensors discussed previously (Section 1.5.1), a blue to red colourimetric transition was observed for these PDA films upon exposure to the appropriate neurotoxin (for G_{M1}-containing biosensor CR = 7% for lactose-PDA; CR = 5% for sialic acid-PDA).[98] Similar results were obtained using the G_{T1b} ganglioside biosensor. The CR was also found to be directly proportional to the quantity of toxin added.[98]

Following on from this work, Charych *et al.*[103] reported the incorporation of G_{M1} gangliosides into PDA liposomes. This was accomplished in an analogous manner to the PDA liposomes which had been developed as influenza sensors (Section 1.5.1) and a similar colourimetric response on exposure to cholera toxin was observed, as for the analogous thin films. In order to prepare the liposomes, G_{M1} (5%) was mixed with the diacetylene lipid 5,7-DCDA as the matrix, rather than the usual 10,12-PCDA. The reason for this difference was because it was found that if 10,12-PCDA (**5**) was employed, the CR of the resulting liposome was significantly reduced.[103] The enhanced sensitivity of the 5,7-DCDA (**4**) system was suggested to be due to the positioning of the diacetylene moiety being nearer to the interface (**4** has three methylene groups instead of eight in **5**). As before, control experiments were carried out to prove that the colour change observed was caused by the specific recognition event and not due to non-specific adsorption.

1.5.3 The use of PDA-based biosensor for the detection of nucleic acids

The development of novel types of DNA-biosensors (such as PDA-based) is largely driven by the needs to effectively diagnose infectious and genetic diseases, detect genetically modified organisms (GMO) and gather forensic evidence. As stated earlier, in order to accomplish this, simple, fast, reproducible and reliable techniques are required. Many groups have reported successful combinations of receptors for the detections of various biological molecules using PDA assemblies. But, from the reports found in the literature it appears that only a few of them have been interested in using PDA-based sensors for the detection of oligonucleotides.

Wang *et al.*[93, 112] have recently reported the use of PDA liposomes functionalised with oligonucleotide headgroups as nucleic acid biosensors. The oligonucleotide probes played the role of sensor and of amplification tag. In their work, Wang and co-workers have prepared two sets of PDA liposomes using a mixture of 10,12-TCDA (**2** - 70%), DMPC (29 mol %) and an oligonucleotides probe modified with cholesteryl (probe 1 or probe 2, 1% - *Figure 36*). The probe was inserted within the liposome surface through hydrophobic interactions. The two probes were prepared with two different sequences of oligonucleotides and were designed to hybridise with only half of the DNA target, each probe targeting a different end of the DNA from the other probe (*Figure 36*). Therefore, 20 μM of target DNA was incubated with a mixture of probes 1 and 2 and the colour of the system turned from deep blue to red within minutes. The corresponding CR was *ca.* 34%. With the decrease of the concentration of target DNA, the colour weakened as well as the corresponding CR decreased. In contrast to this case, the colour transition from blue to red was not observed when target DNA was added to the solutions containing probe 1 or probe 2 only. This showed that the hybridisation of one half of the target did not impose enough mechanical stress to the PDA framework to yield the colour change. When 20 μM of a mismatched DNA target was added to the mixture of probe 1 and 2, no colour changes were observed.



Probe 1 - 5'-*TCTCAACTCGTA*TTTTTT-$(CH_2)_3$-cholesteryl-3';

Probe 2 - 5'-cholesteryl-$(CH_2)_3$-TTTTTT*CGCATTCAGGAT*-3';

Target - 5'-*TACGAGTTGAGAATCCTGAATGCG*-3';

Mismatched - 5'-*GCGTAACTCCTAAGAGTTGAGCTA*-3'

Figure 36 – Schematic diagram of the colourimetric detection of DNA using PDA liposomes functionalised with probe DNA

To address the effect of non-specific interaction between liposomes and target DNA, liposomes composed of DMPC and TCDA only were incubated with 20 μM of target DNA and no colour changes were observed. Although this method demonstrates high specificity for the detection of the DNA target, it has some drawbacks. In fact, the application of this method requires that for each target to be detected, two sets of PDA liposomes, modified with a probe sequence corresponding to half of the target, have to be prepared. Also, it implies the detection of single stranded nucleic acids, otherwise an extra denaturation step of dsDNA would be required and the temperature used for this step might influence the colours of the PDA liposomes (i.e. thermochromism of PDA).

Another approach that uses PDA liposomes for the detection of nucleic acids has been described by Kim and Park.[113] The researchers based the detection on ionic interactions occurring between the positively charged head-groups of PDA liposomes and the negatively charged phosphate backbone of dsDNA. Two different amine-functionalised diacetylene monomers (a primary amine and a quartenary amine) have been designed, prepared and subsequently used to generate PDA liposomes. Therefore, they started with the preparation of the quaternary amine-functionalised PDA liposomes because unlike primary, secondary and tertiary amines, the presence of the positive charge on a quaternary amine is not dependent on the pH of the solution. Thus, the cationic amine-modified diacetylene monomer DADMDPA-*bis*-PDA (DADMDPA stands for 3,3'-diamino-*N*-methyldipropyl-amine) was incorporated in PDA liposomes using 10,12-PCDA (**5**), 10,12-TCDA (**1**) or 2,4-HCDA (**3**) (*Figure 25*, page 31). Due to the difference in the chain length, the resulting liposomes would exhibit their ammonium

headgroups with more or less exposure. The researchers anticipated that the ionic interactions would be maximised if the ammonium groups are fully exposed to the aqueous environment. The difference in exposure is schematically represented in *Figure 37*.

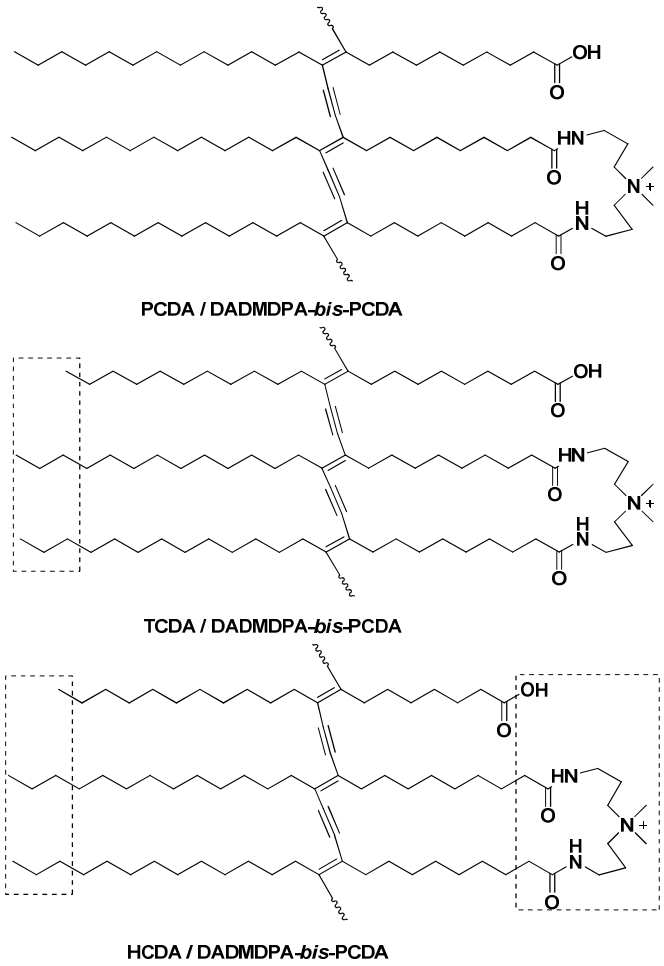
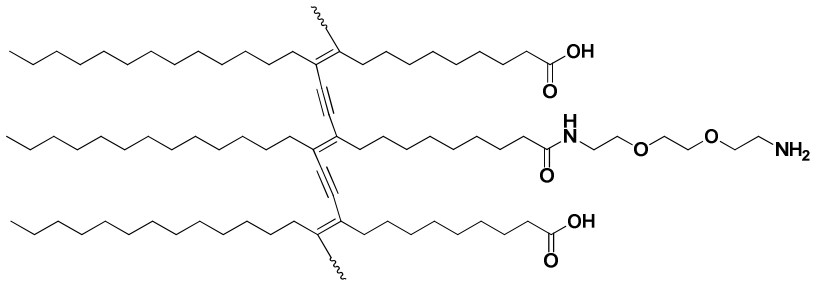


Figure 37

It is clear that, when PCDA is used as matrix lipid, the quaternary amine is not well exposed to the aqueous exterior containing the nucleic acid analyte. Upon the addition of amplified nucleic acids with a size of 605 base pairs (bp) and a concentration of 100 nM no colour changes were observed when the system PCDA/DADMDPA-*bis*-PCDA was tested (CR was 2%).[113] TCDA contains a two carbons shorter tail than PCDA but this small difference did not significantly improve the result as no colour change was observed. A CR of 6.6% was obtained for this system. Finally, HCDA is two carbons shorter in both tail and head part of the molecule. Using the HCDA/DADMDPA-*bis*-PCDA system, the colour of the solution changed from blue to purple and the corresponding CR was 16%. In this system, the quaternary amine was sufficiently exposed to interact with the dsDNA target.

From these results, Kim and Park[113] understood that the full exposure of the amine group is mandatory to induce the colour change. Thus, they prepared a primary amine-modified diacetylene monomers bearing a spacer unit, EDEA [2,2'-(ethylene-dioxy)*bis*(ethylamine)], between the lipid chain and the amine (*Figure 38*). This amine is positively charged in acidic solutions and the liposomes were obtained using the same matrix lipids as earlier (i.e. PCDA, TCDA and HCDA).



PCDA / PCDA-EDEA

Figure 38

Liposomes composed of PCDA/PCDA-EDEA (*Figure 38*) in a molar ratio of 5:5 showed the most dramatic blue-to-red colour change upon the addition of the amplified nucleic acids with a size of 605 bp and a concentration of 100 nM (CR was 46%).[113] Although, the CR obtained was high, this sensing strategy revealed some drawbacks. They relate to the non-specificity of the ionic interactions responsible for the colour change. These positively-charged liposomes have been found to be sensitive towards the PCR reaction buffer solution which contains many salt components. Thus, although, this ionic interaction approach appears limited, it may find use as a rapid test to verify the presence of PCR-amplified dsDNA.

1.5.4 Other applications of PDA-based biosensors

PDA-based liposomes have also been constructed in order to examine enzyme activity. Membrane peptides and membrane bound enzymes (such as phospholipases) have attracted much interest as pharmaceutical drug targets as they are involved in membrane interactions and play major roles in numerous physiological processes (e.g. signalling, cytolysis, ion-channel formation, and cellular recognition). Phospholipases catalyze the hydrolysis of phospholipid molecules at the water-lipid interface of cellular membranes. Phospholipases are involved in diverse biochemical processes, such as fat digestion,

lipid metabolism and the regulation of signal transduction. PLA_2 is an acyl hydrolase which hydrolyses the 2-acyl ester bonds of glycerophospholipids to yield the corresponding fatty acid and lysophospholipid. Phospholipid C (PLC) and phospholipid D (PLD) are phosphodiesterases which attack the phosphate ester moiety within the lipid hydrophilic head-group.

Jelinek and Charych[104] incorporated DMPC into 10,12-TCDA (**1**) liposome. DMPC is a natural substrate of phospholipases so that the activities of three phospholipases were investigated: PLA_2, PLC and PLD.

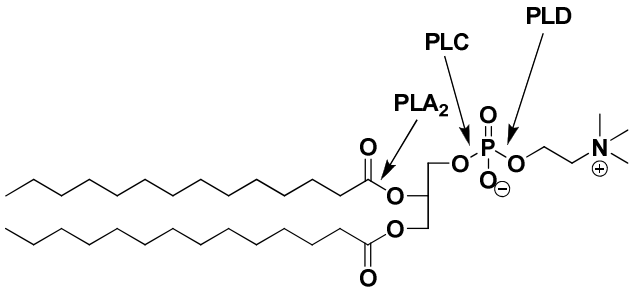


Figure 39 – Structure of DMPC and the points of enzymatic cleavage.

The assemblies in this study was composed of 40% DMPC and 60% of TCDA (*Figure 36*) and following polymerisation, the resulting liposome solution showed a deep blue colour. The solution was treated with each of the phospholipases individually and, in each case, the expected blue-to-red colour transitions were observed (in the case PLA_2, a CR of 15% was obtained). As for the virus and toxin PDA-based biosensors, control experiments were carried out to confirm that it was the specific interaction of the enzymes with the DMPC head-group which was responsible for the observed colour change. In addition, NMR studies were used to confirm the DMPC cleavage sites and to identify for each of the phospholipases the cleavage products.[104] DMPC/PDA liposomes were found responsive to the presence of toxins that exhibit enzyme-like activity, such as β-bungarotoxin (BuTx) isolated from snake venom. This toxin acts at the presynaptic motor nerve termini to alter neurotransmitter release and the distribution of aminophospholipids in the inner- and outer-membrane leaflets.[114] The detection of this toxin is also of significance as it is suspected to be used as potential biological warfare agent. The colour changes of the DMPC/PDA liposomes solutions was observed after only 10 minutes (CR was 25%), although the maximum response was obtained after 50 minutes.[104]

Jelinek *et al.* have also demonstrated that DMPC/PDA liposomes can be used to investigate peptide-membrane interactions.[77] This colourimetric assay has also been coupled with gel chromatography so that the fractions eluted from the separating column are then added to a separate well within a 96-well plate containing DMPC/PDA liposome solution.[78] These PDA liposomes have also shown to afford a colourimetric response in the presence of specific cations, e.g. the capability to distinguish between Na^+ and K^+ has been demonstrated.[94] By adding a hydrophobic peptide displaying an epitope, which recognizes a specific antibody, into the same kind of assembly, Jelinek and Kolusheva showed that it was possible to investigate specific antibody-epitope interactions.[76]

Jiang *et al.* reported the colourimetric detection of the bacteria *E. coli* using glucoside decorated PDA liposomes.[92] These liposomes were prepared, separately, from 2,4-TCDA and 10,12-PCDA with 2% dioctadecyl glycerylether-β-glucoside (DGG) incorporated (*Figure 40*).

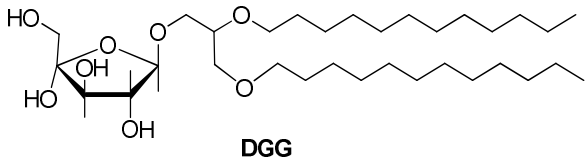


Figure 40

E.Coli was dispersed in an aqueous solution of sodium chloride and this was then added to the PDA liposome solutions. In the case of liposomes made from PCDA, no colour changes were observed, whereas with the liposomes made from TCDA, the system turned from deep blue to red within seconds. No colour changes were observed when the solution of sodium chloride only was added as well as when liposomes containing no DGG were stirred with the solution of *E. coli.* These tests showed that DGG/*E.Coli* interactions are responsible for the colour transition of the PDA liposomes.

Li *et al.* have also described colourimetric detection of *E. coli*. In this case, 10,12-PCDA liposomes were constructed with 5% mannoside conjugated to a hexadecane tail embedded.[115] The colourimetric response of the system was 62%. In order to show specificity of the detection system, *E. coli* was first mixed with TiO_2 colloids which acted as a photocatalyst. UV irradiation of this solution for 30 minutes leads to the de-

struction *E. coli*, which was evidence by the lower value of the colourimetric response of the liposomes upon addition of this latter solution (CR 10%).

Cheng and Stevens[116] synthesised a series of amino acid-modified 10,12-PCDA lipids and incorporated them into PDA liposomes (*Figure 41*). ILE-PDA and PHE-PDA lipids, both containing hydrophobic segments in the head-group region, did not produce liposomes, showing that the head-group requires hydrophilic character in order to align and self-assemble.

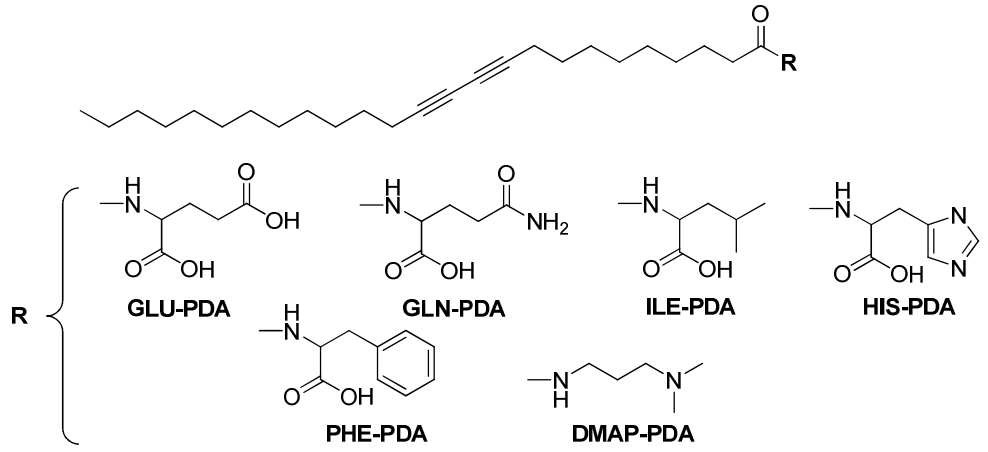


Figure 41

The liposome solutions using lipids GLU-PDA, GLN-PDA and HIS-PDA respond immediately to an increase of pH by undergoing rapid blue to red colour transitions.[116] HIS-PDA also showed some unique properties in response to acidic pH change, while GLU-PDA and GLN-PDA did not. This can be attributed to the protonation of the imidazole ring in the HIS-PDA lipid. The same kind of response was obtained with amine-terminated DMAP-PDA lipid. It was suggested that the colour change was due to the charge-induced head-group rearrangement causing a conformational change in the backbone and hence inducing a colour change.[116]

In conclusion, the variety of shapes obtained for self assembling PDA-based system as well as the wide variety of binding interactions described in this section indicates the potential of PDA-based systems as effective colourimetric biosensors. In particular, the ability that the PDA-liposomes have to mimic cellular membrane environment and shape leads to a variety of novel biochemical applications.

The use of PDA does not require sophisticated instrumentation for analysis and could be incorporated into portable devices. The colourimetric system described is robust and versatile; the liposome solutions can be stored at 4 ^{o}C and retain their colour sensitivity for long periods of time (i.e. months). The colour response is rapid and with the incorporation of the correct head-group, the response can be highly specific. Although the majority of the examples described previously showed that the colour response is irreversible, new developments into this field have shown the possibility for the colour change to be reversible, leading to reusable bioassays.

Finally, it is worth noting that the recognition element can be either covalently linked to the PDA framework through copolymerisation or embedded within the lipid bilayer by hydrophobic interactions. The latter, therefore, allows the incorporation of natural molecules into synthetic liposomes and avoids difficult synthetic routes.

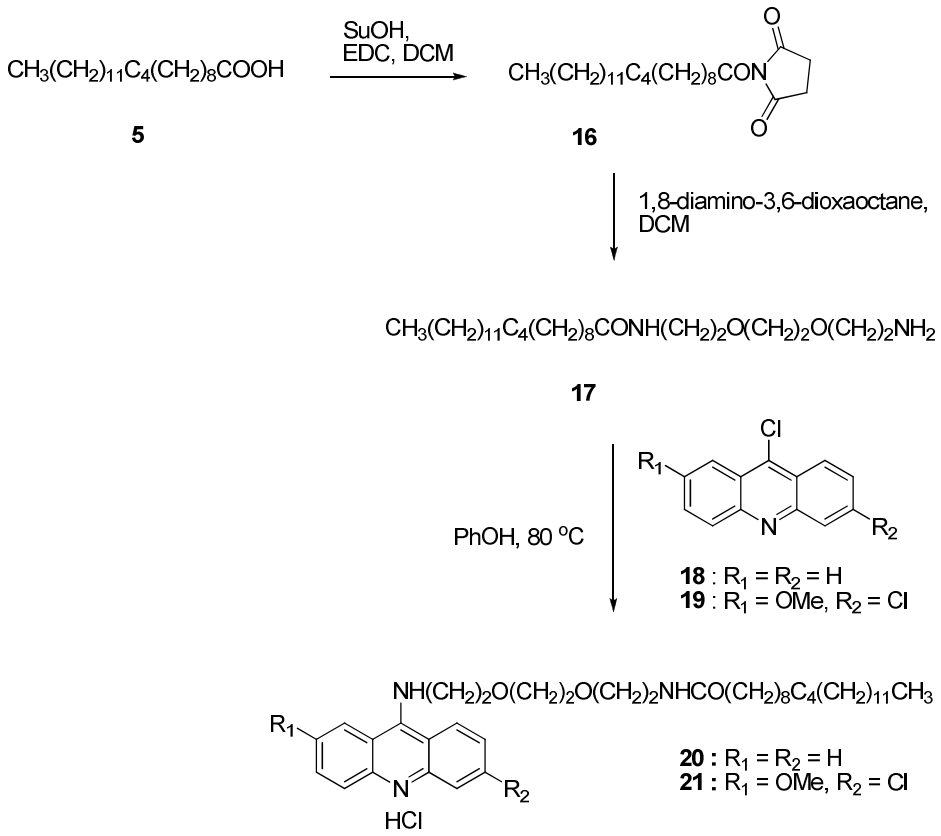
1.6 PREVIOUS WORK IN THE GROUP INVOLVING PDA

Prior to embarking on the research described in this thesis, two other projects have been undertaken at Heriot-Watt University which involved the preparation and investigation of suitably functionalised PDA liposomes for use as a nucleic acid biosensor.

1.6.1 Acridine-functionalised PDA liposomes

Murray[117] reported the design of a novel nucleic acid binding assay based on the chromic effects of acridine-functionalised PDA liposomes. In order to prepare these liposomes, the two acridine derivatives (**20**) and (**21**) containing a diacetylenic lipid tail, an ethylene glycol spacer and the acridine intercalator as the head-group were first synthesised. This was successfully achieved following the route outlined in *Scheme 2*.

Commercially available 10,12-PCDA (**5**) was converted into *N*-succinimidyl-10,12-pentacosadiynate (**16**) in 93%. The desired amine **17** was then obtained by dropwise addition of a solution of **16** through a syringe pump to a solution of the co-reactant diamine over 16 hours (38%). Subsequently, the target acridines **20** (74%) and **21** (69%) were prepared by heating **18** and, separately, **19** with the amine **17** in a solution of phenol.[118]



Scheme 2

PDA liposomes were subsequently constructed containing 5% of models **20** or **21** and 95% of the matrix lipid, *N*-(8-amino-3,6-dioxaoctyl)-10,12-pentacosadiynamide (**17**). Murray[117] discovered that the protocol reported by Charych *et al.*[90] for the preparation of PDA liposomes involving sonication of the lipid mixture at 70-80 °C could not be utilised in this case. The reason for this was because matrix lipid **17** spontaneously polymerised over this range of temperatures. The liposome solutions were thus prepared at room temperature. Subsequently, polymerisation of the acridine-containing liposomes solutions **20** and **21** was achieved using an insertion apparatus and a 400 W medium pressure Hg lamp. This gave deep blue polymerised liposomes with electronic absorption spectra showing λ_{max} at 648 nm and 651 nm, respectively. The resulting blue solutions containing either **20** or **21** were tested for their abilities to exhibit a colourimetric response in the presence of salmon testes DNA (ST-DNA). Thus, a solution of ST-DNA was added to the cooled liposome solutions and the mixtures were left to stir for 48 hours. However, after this time, complete precipitation of the liposomes made from **20** and **21** had occurred and there was no evidence of a colour change (no visible changes or modifications in their electronic absorption spectra). This research on the acridine-functionalised PDA liposomes was not taken any further because it was rea-

soned that such liposomes would have only a limited application as nucleic acid biosensors since acridine is a sequence non-specific DNA intercalator.

1.6.2 Peptide nucleic acid functionalised PDA liposomes

Staying with the aim to prepare a novel colourimetric nucleic acid biosensor based on suitably functionalised PDA liposomes, the group at Heriot-Watt University revised their initial idea to use acridine type compound as the recognition element based on the limitations reported above. Indeed, they decided to explore the use of peptide nucleic acids as the head-group. As stated previously (Section 1.3) PNAs offer a number of advantages over natural oligonucleotides regarding their use in biosensors; namely they show higher selectivity and stability.

Thus, Murray[2] developed synthetic routes for the preparation of model compounds composed of a PNA monomer head-group connected *via* either its *C*- or *N*-terminus to a saturated stearoyl or an unsaturated diacetylenic tail, (models **22**, **23** and **24** – *Figure 42*). It was necessary to first prepare and investigate these monomer model compounds in order to determine whether they could be successfully incorporated into PDA liposomes and that the PNA head-group did not hinder the required topochemical polymerisation from taking place. In fact Charych *et al.*[68] have reported that the prediction of successful polymerisation is not possible.

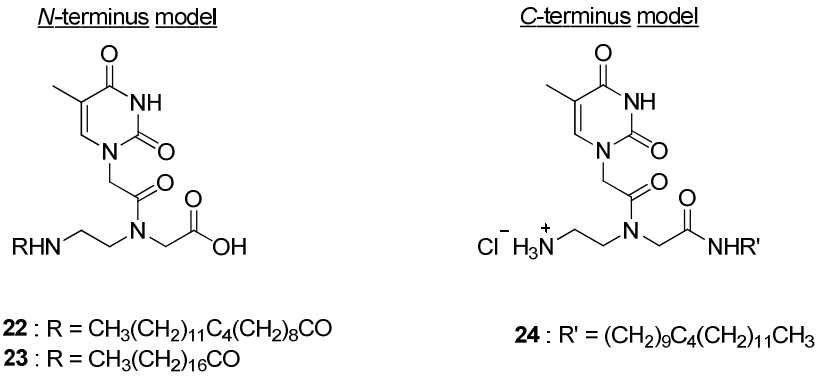
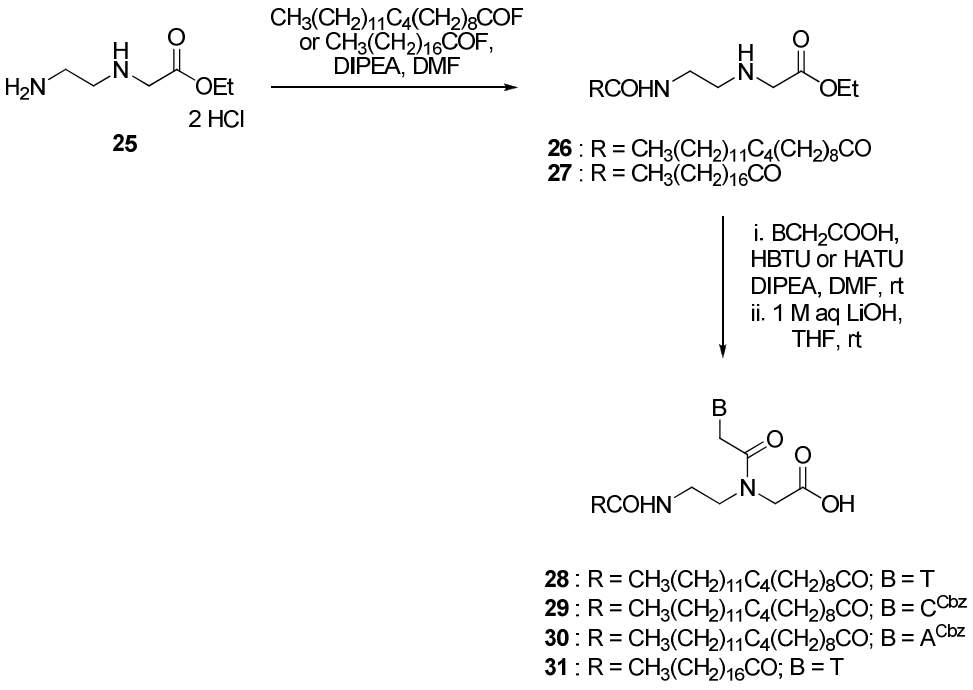


Figure 42

Figure 42 shows only the models developed bearing thyminyl PNA monomer but the analogous *N*-terminus models bearing the adeninyl and the cytosinyl PNA monomer were also synthesised. The preparation of the corresponding PNA monomers has been extensively reported in the literature[1] and will be briefly described in Chapter 2.

Murray[2] started his investigations with the preparation of the *C*-terminus monomer model compounds such as **24** (*Figure 42*). Subsequently, **24** was successfully incorporated into liposomes using the matrix lipid previously employed for the acridine-functionalised PDAs (**17** – *Scheme 2*). The solution obtained was deep blue. Unfortunately, as reported earlier by Murray[117], the liposomes prepared precipitated very quickly at room temperature. Therefore, due to these preliminary results the *C*-terminus strategy was abandoned.

Consequently, Murray[2] investigated the preparation of *N*-terminus model compounds such as **22** and **23** (*Figure 42*). The design of the best synthetic route was inspired by literature precedent[119] and involved the preparation of the PNA backbone unit-functionalised with the desired lipid tail (**26** or **27** - *Scheme 3*). **26** (65%) and **27** (60%) were obtained from the coupling between the PNA backbone **25** and the acid fluoride derivative of the chosen lipid tail. The free amino group of both **26** and **27** were subsequently coupled with the appropriate nucleobase acetic acid derivative (thyminyl, cytosinyl or adeninyl) using HBTU as coupling agent. Models **28**, **29**, **30** and **31** were obtained in 97%, 95%, 69% and 97% yields, respectively.



Scheme 3

Due to time constraints, Murray[2] incorporated only the thyminyl models **28** and **31** into PDA liposomes. Blue solutions were observed upon polymerisation of mixtures con-

taining 5% of the acid models **4d** or **4g** with 95% of 10,12-PCDA (**5** – *Scheme 2*) matrix lipid. No differences between the saturated and diacetylenic model were detected visually and in the electronic absorption spectra. These results suggested that the presence of a PNA monomer did not hinder the formation of PDA liposomes or the polymerisation reaction. Murray[2] discovered that the presence of a free terminal carboxylic acid was mandatory in order to avoid precipitation of the liposomes from the aqueous solution that had been observed with the corresponding ester models.

Murray[2] then tested the ability of these models to undergo colourimetric changes. Thus, a solution of adenosine in PBS buffer pH 7.4 was added to the liposome solutions containing separately the saturated model **31** and the diacetylenic model **28**. Unfortunately, no colour changes were recorded. This implied that the formation of the expected T-A complex did not result in enough mechanical stress to induce the 'blue-to-red' colour transition, probably because the Watson-Crick hydrogen bonding within a single base pair are not strong enough. It is also possible that the T-A duplexes have not been formed.

Based on this finding Murray[2] decided to investigate the preparation of PDA liposomes bearing PNA oligomers. It was envisaged that the PNA:DNA heteroduplex formation would induce more mechanical stress to the PDA framework and thus afford the desired colour changes. Therefore, Murray[2] synthesised PNA $(T)_{10}$ oligomers manually in a stepwise fashion on a suitably derivatized solid support employing the solid-phase protocol described by Christensen *et al.*.[120]

32 : $CH_3(CH_2)_{16}CO\text{-}(T)_{10}\text{-}LysNH_2$

33 : $CH_3(CH_2)_{16}CO\text{-}(T)_{10}\text{-}AspNH_2$

34 : $Ac\text{-}(T)_{10}\text{-}Lys\{CO(CH_2)_{16}CH_3\}NH_2$

35 : $Ac\text{-}(T)_{10}\text{-}Lys\{CO(CH_2)_8C_4(CH_2)_{11}\}NH_2$

36 : $AcAsp\text{-}(T)_{10}\text{-}Lys\{CO(CH_2)_{16}CH_3\}NH_2$

37 : $AcAsp\text{-}(T)_{10}\text{-}Lys\{ CO(CH_2)_8C_4(CH_2)_{11}\}NH_2$

38 : $CH_3(CH_2)_{11}C_4(CH_2)_8CO\text{-}(T)_{10}\text{-}AspNH_2$

Figure 43

The stearoyl group was incorporated at the *N*-terminus of oligomers **32** and **33** (*Figure 43*) in the final coupling step once the PNA $(T)_{10}$ had been assembled on the solid sup-

port using the stearoyl fluoride derivative. The same approach using 10,12-pentacosadiynoyl fluoride was unsuccessful. Therefore, in order to prepare models **33**, **34**, **37** and **38** Murray[2], successfully, used a post-synthetic solution-phase approach in which he reacted the cleaved the PNA-$(T)_{10}$ oligomers with 10,12-pentacosadiynoyl fluoride in the presence of diisopropylethylamine (DIPEA).

The formation of liposomes from two component mixtures of acid terminated models **33**, **36**, **37** and **38** (5%) and 10,12-pentacosadiynoic acid (**5** – *Scheme 2*) were subsequently investigated using a procedure similar to that described by Charych *et al.*.[98, 103] Deep blue solutions of liposomes were obtained from samples containing diacetylenic-PNA lipids **36**, **37** and **38** or the stearoyl-PNA lipid **33**. The solutions showed some precipitation after a couple of days at room temperature. The mixture of **36** (5%) and **5** (95%) gave only a dilute, light blue solution of liposome and much precipitation occurred during the preparation.

Due to time constraints, Murray[2] only managed preliminary studies of the abilities of these models to exhibit a colourimetric response upon hybridisation to the complementary PNAs. H-A_{10}Lys-NH_2 PNA oligomer in PBS buffer at pH 7.4 was therefore added to the liposome solutions, under similar conditions that have been used by Charych *et al.*.[91, 121, 122] Unfortunately, no colour transitions were recorded. It was rationalised that the precipitation of the liposomes during these studies was at the origin of this failure. Nevertheless, Murray's results[2] showed that it was possible to incorporate PNA oligomers into PDA liposomes and that the solutions obtained turned blue upon photopolymerisation, therefore, the presence of the PNA oligomer as the head-group did not hinder the formation of liposomes. He also showed that the PNAs could be functionalised with either saturated or unsaturated (diacetylene-containing) lipids and that the electronic absorption spectra of both solutions were comparable. The colourimetric studies were only preliminary and further work into the field should demonstrate that PNA-functionalised PDA liposomes can be used as colourimetric nucleic acid biosensors.

CHAPTER 2: SYNTHESIS OF MODELS

2 SYNTHESIS OF MODELS

2.1 AIMS AND OBJECTIVES

As stated in Section 1.6.2 (page 53), the ultimate aim of this research was to develop a novel, rapid and label-free DNA detection assay employing a colourimetric biosensor derived from PNA-functionalised PDA liposomes. The overall objectives of the work reported in this thesis was to synthesize and evaluate lipid-functionalised PNA monomers bearing a more favourable hydrophilic/lipophilic contrast than the models reported by Murray *et al.*[2] so that the PNA headgroups were sufficiently hydrated under the conditions required for construction of the liposomes.

Schneider *et al.*[123, 124] have extensively studied peptide nucleic acid amphiphiles (PNAA). These conjugates are comprised of PNA molecules covalently linked to one or more *n*-alkane chains (*Figure 44*).

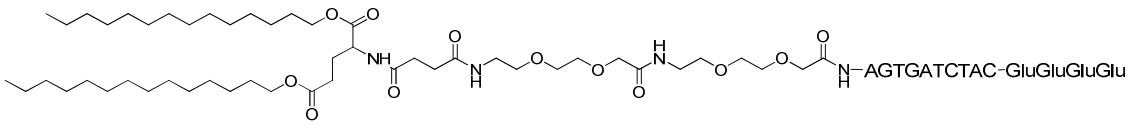


Figure 44 – PNAA: $(C_{14})_2$-$(AEEA)_2$-AGTGATCTAC-$(Glu)_4$

Since PNA is uncharged, it is sparingly soluble in water and it has a tendency to self-aggregate in solution. Nielsen *et al.*[45] have reported that the inclusion of a terminal lysine group greatly improves their solubility. In PNAA, Schneider *et al.*[123, 124] have reasoned that the covalent attachment of one (or more) alkane chains would exacerbate the poor water solubility of PNA. By attaching charged amino acids (e.g. glutamic acid and lysine) to the PNA head-group, these researchers have imparted sufficient hydrophobic/hydrophilic contrast for micellization and water solubility of the corresponding PNAA.[123] As a rule, they reported that four charged amino acids per 10-mers (or two per 6-mers) are required for adequate water solubility. They also demonstrated that PNA peptide can absorb on the surface of the liposome resulting in poor hybridisation with complementary DNA and this limitation was overcome by introducing ethylene glycol spacers, such as 8-amino-3,6-dioxaoctanoic acid (AEEA) (*Figure 44*).[124, 125]

Charych *et al.*[90] have also included a similar spacer (i.e. AEEA) between the sialic acid head-group and the diacetylene reporter element in their influenza virus biosensors (Section 1.5.1, page 37). They reasoned that this spacer extended the bioactive head-group beyond the matrix lipid and allowed a better contact with the virus.

In the light of Schneider's and Charych's findings, we decided to investigate whether the inclusion of a similar AEEA spacer could be successfully incorporated into the PNA-functionalised PDA liposomes developed by Murray *et al.*[2] and improve their water solubility. It was envisaged that an ethylene glycol spacer could be covalently linked to the *N*-terminus of PNA at one end and the lipid tail at the other through amide bonds, using established peptide chemistry methods (*Figure 45*).

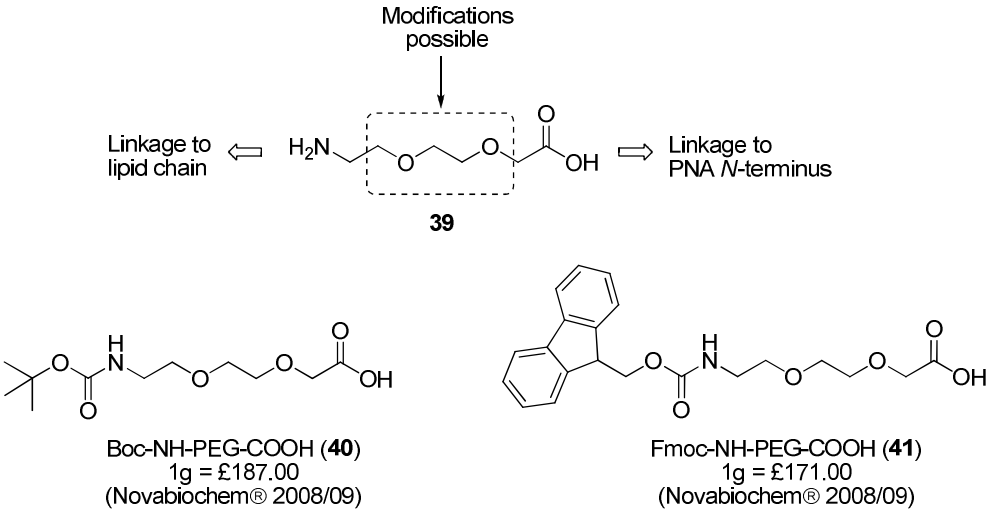


Figure 45

Due to the cost of *N*-protected derivatives of AEEA (**40** and **41** - *Figure 45*), we first required a viable synthetic route to this compound in order to use it for our purpose. We reasoned that developing our own pathway to AEEA would also allow us to explore the utility of a range of "PEG-like" spacer analogues (e.g. the effect of increasing the CH_2CH_2O repeating unit) for constructing PNA-functionalised PDA liposomes.

Subsequently, following the development of a suitable route to AEEA spacer **39** (*Figure 45),* we planned to study the inclusion of this moiety into model 'lipid-spacer-PNA' conjugates (*Figure 46*) bearing either thyminyl, cytosinyl and adeninyl PNA monomers. We envisaged that once routes to these model conjugates had been optimised, they could be simply extended to the synthesis of conjugates containing PNA oligomers.

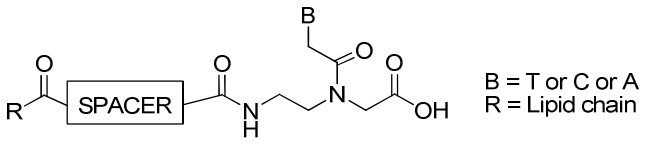


Figure 46

Providing that these 'lipid-PEG-like spacer-PNA' conjugates could be prepared, we proposed to subsequently explore the ability of these compounds to be incorporated into PDA liposomes, using 10,12-PCDA (**5**) as the matrix lipid. If coloured solutions are obtained from these model conjugates, we planned to study the electronic absorption spectra of their 'blue' and 'red' forms. Following these studies, we hoped to undertake preliminary investigations into their stability, and finally, to study whether any colourimetric response could be obtained upon their exposure to complementary nucleic acids.

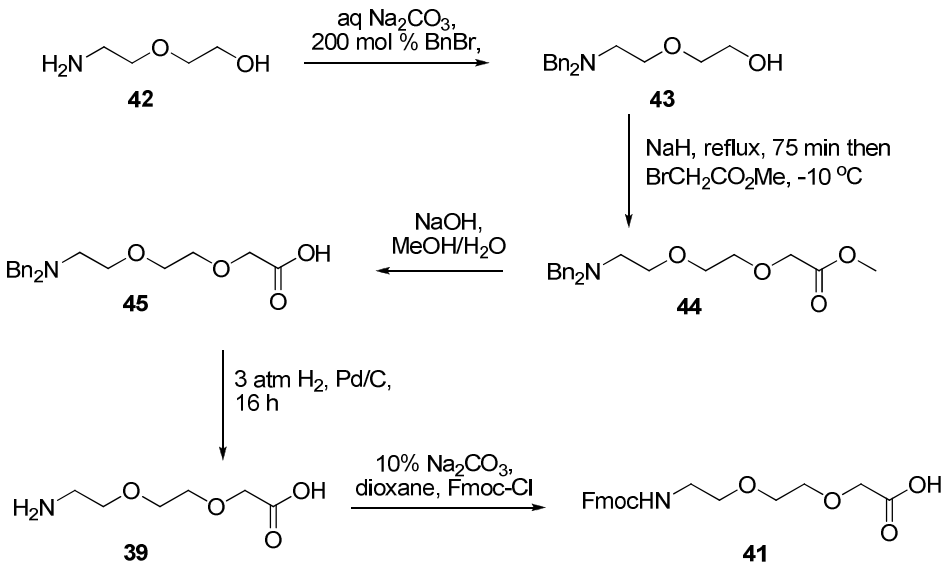
Our progress towards achieving these aims and objectives will now be discussed in the remainder of Chapter 2 and in Chapter 3.

2.2 SYNTHESIS OF 'LIPID-PEG-LIKE SPACER-PNA' MODELS

2.2.1 Preparation of the 'PEG-like' spacer

The preparation of PNA-functionalised PDA models bearing a "PEG-like" spacer required first development of a viable synthetic route to the Boc-AEEA spacer **40** (*Figure 45*). We chose to have the amino terminus Boc-protected as this protecting group would be compatible with our established in-house solid phase strategy for preparation of PNA oligomers. We envisaged that, at a later stage in this research, it should be possible to attach spacer **40** to the *N*-terminus of the PNA oligomer as part of this solid phase approach.

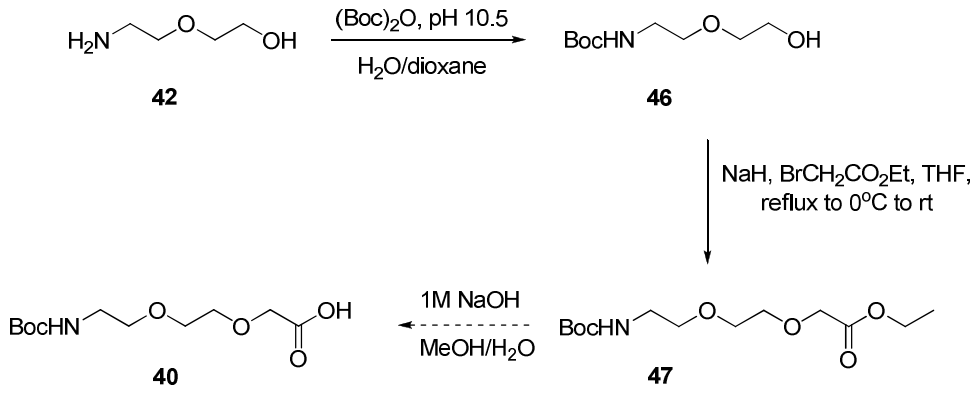
Koskinen *et al.*[126] have recently reported the preparation of *N*-Fmoc-8-amino-3,6-dioxaoctanoic acid **41** starting from commercially available 2-(2-aminoethoxy)ethanol **42** as outlined *Scheme 4*.



Scheme 4

In their route, the amino function of **42** was first protected with two benzyl groups by reaction with benzyl bromide in the presence of sodium carbonate at room temperature. The dibenzyl analogue **43** was obtained in 68% yield. Subsequently, alkylation of the hydroxyl group of **43** was achieved with methyl bromoacetate in the presence of sodium hydride to afford ester **44** in 96% yield. Ester **44** was then hydrolysed using aqueous sodium hydroxide to afford the carboxylic acid **45** in 80% yield. Consequently, 8-amino-3,6-dioxaoctanoic acid (**39**, 91%) was obtained by hydrogenolysis of **45**. Finally, the amino function of amino acid **39** was protected with an Fmoc group to give the desired compound **41** in 69% yield.

As mentioned earlier, we were interested in preparing the Boc analogue **40** (*Scheme 5*) because this protecting group was compatible with our established Boc solid phase protocol used by Murray *et al.*[2] for the preparation of PNA oligomers. However, we reasoned that we could adapt the route developed by Koskinen *et al.* to yield the required Boc-protected compound. Our revised route is outlined in *Scheme 5*. In addition to requiring the use of a Boc protecting group, we also needed to avoid the hydrogenation step used by Koskinen *et al.* to remove the benzyl groups. This was because the Department of Chemistry at Heriot-Watt University did not have a functioning Parr hydrogenation apparatus, needed for the reaction, at the time of this research.



Scheme 5

Thus, the first step in our synthetic pathway involved the protection of the amino moiety of **42** with a Boc group (*Scheme 5*). This was sucessfully achieved by treatment of a solution of **42** in a mixture of water and dioxane (10:1 v/v) with di-*tert*-butyl dicarbonate. Following work-up and purification by column chromatography, compound **46** was obtained in 90% yield. The appearance of a singlet (integrating to 9H) at 1.34 ppm in the ^{1}H NMR spectrum recorded for **46** and the two extra peaks at 28.3 ppm and 79.2 ppm in ^{13}C NMR confirmed the introduction of the *N*-Boc protecting group.

The next step in our route involved the alkylation of the hydroxyl function of **46** to yield the ethyl ester **47**. We chose to follow the protocol described by Koskinen *et al.*.[126] This involved the slow addition of one equivalent of sodium hydride to a solution of **46** in anhydrous THF followed by heating the resulting solution to reflux for 75 minutes. After this time, the solution was cooled to 0 ^{o}C and ethyl bromoacetate was added dropwise. After work-up, a very complex mixture of products was afforded. And after careful purification by column chromatography, ester **47** was isolated in 10 % yield. Another product was also recovered from this mixture, which was shown to be unreacted starting material **46** from ^{1}H and ^{13}C NMR spectroscopy. The reason for the low yield of ester **47** was not apparent. However, we had concerns about leaving the reaction mixture for such a long period at high temperature after the addition of sodium hydride and before the addition of the alkyl halide. Alkoxide anions are highly moisture sensitive and particular care must be taken to ensure complete dryness of the reaction medium. Temperature is also known to increase the probability of side reactions from occurring. Therefore, we reasoned that the alkoxide of **46**, once formed, would be highly reactive and would have a short half-life. Thus, in order to reduce the risk of side reactions occurring, we decided to repeat the reaction at room temperature. Unfortunately, no im-

provement in the yield of **47** was observed (10%). We then decided to explore the effect of the solvent on the reaction. Polar aprotic solvents such as DMSO, DMF, acetonitrile and THF have the ability to stabilize the alkoxide anion formed and therefore are favorable for S_N2 reactions. Since DMF and acetonitrile are more polar than THF, they should therefore improve the reaction. Thus, we decided to repeat the reaction using DMF instead of THF. In this case a slight improvement in the yield of **47** was obtained (19%) but this yield was clearly still inadequate for our purpose.

At this point it was unclear whether the alkoxide anion of **46** had formed and that, due to its poor stability, it accepted a proton to reform the starting material or if, under these reaction conditions, the hydroxyl proton of **46** had not been properly abstracted. It has been reported that the presence of electron-withdrawing groups attached to the alcohol group will increase the acidity of the -OH proton while the presence of electron-donating groups will decrease it. In the case of the later, the alkoxide anion formed will be less stable and so will tend to readily accept a proton to reform the original alcohol. Thus, in our case, we hypothesize that the presence of the alkyl group next to the alcohol did not help to stabilize the corresponding alkoxide anion. In addition to the stability of the resulting alkoxide of **46**, we should recognize that this compound has two potentially acidic protons: the proton of the carbamate group (in red) and the proton of the alcohol (in blue - *Figure 47*). This was not a concern for Koskinen *et al.* as they had worked with the dibenzyl protected analogue **43** (*Scheme 4*) rather than the Boc-protected compound **46**. The exact pK_a values of these protons in **46** are not known but from the data reported by Bordwell *et al.*[127] we were able to estimate their values (*Figure 47*).

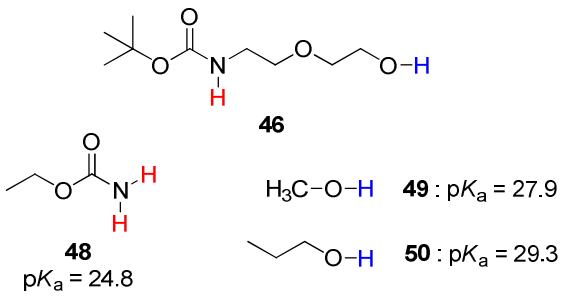


Figure 47 – pKa values measured in DMSO[127]

The pK_a value of the proton of carbamate **48** is 24.8, while the pK_a values of methanol (**49**) and *iso*-propanol (**50**) are 27.9 and 29.3, respectively. This means that the proton of

a carbamate **48** is *ca.* 1000 times more acidic and therefore more easily deprotonated than the hydroxyl proton of either methanol (**49**) or *iso*propanol (**50**). This finding might explain why, in our case, the yield of the ester **46** was so low. Unfortunately, we did not isolate any of the possible *N*-alkylated side product (**51** – *Figure 48*) from the complex mixture of products. Therefore, we were unable to confirm whether this is an issue.

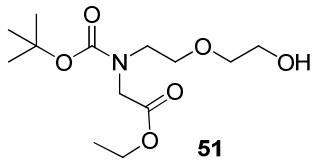
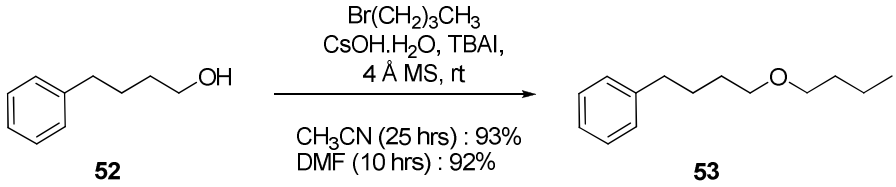


Figure 48

In an attempt to see whether this alkylation could be achieved under alternative conditions, we decided to investigate other bases in this reaction. Jung *et al.*[128] have reported the use of cesium hydroxide monohydrate for the preparation of a series of ethers from primary alcohols. Cesium bases have shown a number of advantages in alkylation reactions with one being that the *in situ* generated cesium alkoxides are considered to be more nucleophilic than their corresponding alkaline earth metal alkoxides due to their weak solvation.[129] For example, Jung *et al.*[128] treated a solution of alcohol **52** (*Scheme 6*) with bromobutane in the presence of cesium hydroxide monohydrate, TBAI and molecular sieves.

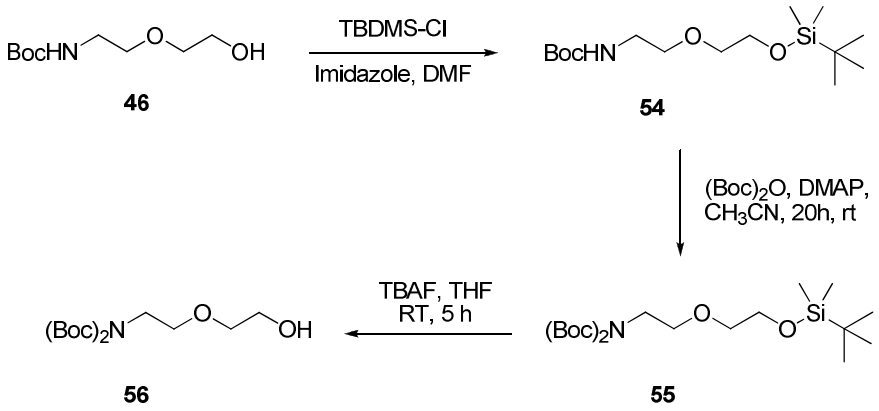


Scheme 6

The corresponding ether **53** was subsequently obtained in yields of more than 90% when either anhydrous acetonitrile or anhydrous DMF was the reaction solvent. The presence of powdered 4 Å molecular sieves was used to remove traces of water and therefore prevent the alkoxide, once generated, from being protonated. The contribution from TBAI stemmed presumably from a Finkelstein type reaction and phase transfer catalysis.[128]

Therefore, the conditions described by Jung *et al.*[128] seemed very promising for our purpose. We decided to first investigate the use of one equivalent of cesium hydroxide

monohydrate without the contribution of TBAI in our initial study. Thus, a solution of alcohol **46** (*Scheme 5*) in anhydrous DMF was treated with cesium hydroxide monohydrate in the presence of 4 Å powdered molecular sieves for 30 minutes at room temperature before one equivalent of ethyl bromoacetate was added. After work-up and purification by column chromatography, none of the products isolated from the complex reaction mixture corresponded to the expected alkylated product **47**. Subsequently, we decided to repeat the reaction in the presence of one equivalent of TBAI and this time product **47** was obtained in an 18% yield. Since Jung *et al.* reported similar a yield when anhydrous acetonitrile was used instead of anhydrous DMF, we decided to repeat the reaction using anhydrous acetonitrile. However, in our case the yield for **47** was lower than in anhydrous DMF (10%). We also repeated the reaction using potassium carbonate and cesium carbonate in anhydrous DMF but in both cases, ester **47** was isolated in a low yield of 17%. Due to our concerns regarding selective deprotonation of the alcohol function relative to the carbamate proton, we decided to see if this problem could be overcome by the use of the di-Boc protected analogue **56** (*Scheme 7*). After all, Koskinen *et al.*[126] had selected to protect the amino function of **42** with two benzyl groups.



46 **54** **56** **55**

Scheme 7

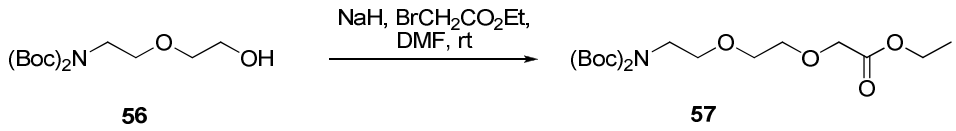
A route to the di-Boc derivative **56** was therefore required starting from mono-Boc protected alcohol **46**. We chose to start from **46** rather than the free amino alcohol **42** because **46** was readily available in the lab in large amounts. We decided to employ the procedure described by Appella *et al.*[130] for the preparation of the di-Boc compound **56**. The conditions used by this group to prepare their di-Boc analogues implied the treatment of a solution of a mono-Boc protected amino acid in anhydrous acetonitrile with an excess of di-*tert*-butyl dicarbonate in the presence of 4-dimethylaminopyridine (DMAP). DMAP is also known to react preferably with alcohols than with amino function, therefore, this function in **46** required to be protected before the second Boc group

could be inserted. Thus, we selected to use the *tert*-butyldimethylsilyl protecting group for this purpose. A solution of alcohol **46** in anhydrous DMF was therefore treated with *tert*-butylchlorodimethylsilane and imidazole. After work-up and purification by flash column chromatography, silyl ether **54** was obtained as a white solid in 84% yield. In the ^{1}H NMR spectrum recorded for **50**, a singlet appeared at 0.05 ppm which was attributed to the two methyl groups attached to the silicon atom. The singlet at 0.87 ppm corresponded to the nine *tert*-butyl silyl protons.

Subsequently, the di-Boc protected analogue **55** was obtained by treating a solution of **54** in anhydrous acetonitrile under the same conditions as described earlier by Appella *et al.*.[130] The progress of the reaction was monitored by TLC. Following completion of the reaction, work-up and purification by column chromatography, the expected product **55** was obtained in 80% yield. In the ^{1}H NMR spectrum recorded for **55**, the singlet at 1.48 ppm integrated now for 18 protons rather than nine as in the analogous spectrum recorded for **54**. This confirmed the presence of two Boc groups attached to the same atom in **51**. HRMS analysis further confirmed the molecular formula of the product obtained.

The final step in the synthetic pathway involved the cleavage of the silyl ether protecting group from **55**. This was accomplished under standard conditions using TBAF. After leaving a solution of **55** and TBAF in THF to stir at room temperature for five hours, TLC analysis showed that no starting material remained. Work-up and flash column chromatography gave the expected product **56** in a quantitative yield. The disappearance of the proton signals corresponding to the silyl ether moiety in the ^{1}H NMR spectrum together with the appearance of a broad singlet at 2.47 ppm confirmed that the removal of the silyl ether had been successful and that the corresponding alcohol **56** has been formed.

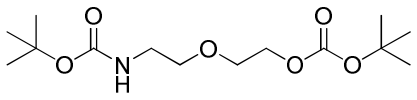
Having successfully prepared the required di-Boc protected alcohol **56**, we then embarked upon investigation of its ability to be *O*-alkylated (*Scheme 8*).



56 **57**

Scheme 8

Our initial attempt involved treatment of a solution of alcohol **56** in THF with sodium hydride followed ethyl bromoacetate under the same conditions as those described by Koskinen *et al.*.[126] After work-up and flash column chromatography, spectral analysis of the fractions isolated failed to show the presence of the desired ether **57** (*Scheme 8*). However, a side reaction product was isolated from the complex crude reaction mixture which was identified as carbonate **58** (40% - *Figure 49*).



58

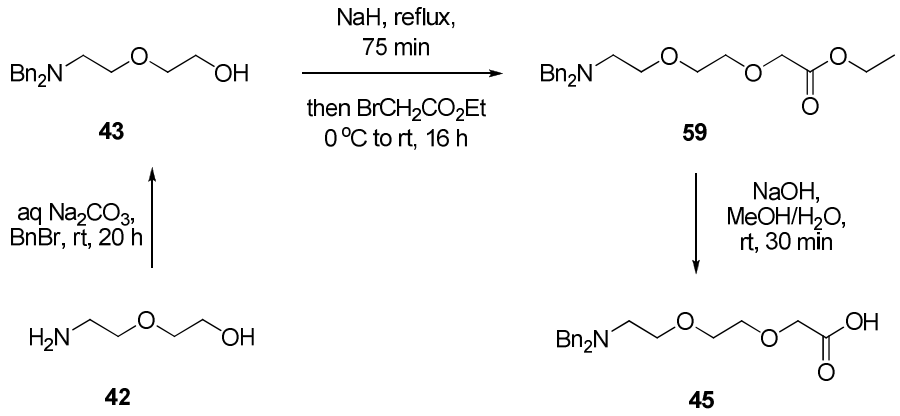
Figure 49

Despite this result, we decided to continue our examination of this reaction. Our next attempt involved a treatment of a solution of **52** and ethyl bromoacetate (3 eq.) in anhydrous THF, cooled to 0 ^{o}C, with sodium hydride (1.5 eq.). By performing the reaction in this manner, it was hoped that the ethyl bromoacetate would react instantly with the alkoxide anion of **56** as soon as it was formed. This reaction met with some success with **57** being isolated in a 19% yield, following work-up and purification by column chromatography. However, the yield of **57** was similar to those obtained earlier for the alkylation of the mono Boc derivative **46**. Surprisingly, carbonate **58** had been produced again, this time in a 68% yield. Unfortunately, the reason for this is unknown.

At this time in the project, we realised that the preparation of the *N*-Boc protection version of the desired AEEA spacer was going to be more challenging than first expected. Since time was moving on, we decided to revise our initial strategy for preparation of the Boc-AEEA spacer **40**. We reasoned that, although the Boc protected compound was required, it did not have to be introduced at the beginning of the scheme. Instead, the Boc protecting group could be introduced after 8-amino-3,6-dioxaoctanoic acid (**39** – *Scheme 4*) had been prepared in the same way that Koskinen *et al.*[126] had prepared the Fmoc *N*-protected analogue **41**. Therefore, we decided to follow Koskinen's approach exactly, initially protecting the amino function of the starting material **42** with a diben-

zyl protecting group. This would subsequently be cleaved and replaced in the last step of the synthetic pathway for our desired Boc protecting group.

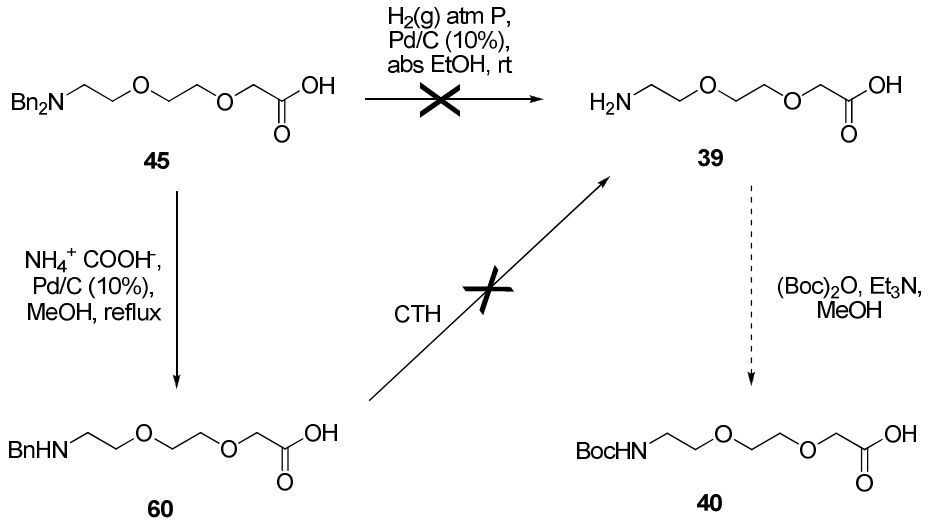
The first step in this strategy concerned the protection of the amino moiety of **42** with two benzyl groups. The reaction was carried out in the same manner as described by Koskinen *et al.*[126] (page 61) and dibenzylated product **43** was obtained in 80% yield. This yield was comparable to that reported by Koskinen (*Scheme 9*). The next step in this pathway involved *O*-alkylation. Thus, a solution of alcohol **43** in anhydrous THF was heated to reflux in the presence of sodium hydride for 75 min. After this time, the solution was cooled to room temperature and ethyl bromoacetate was slowly added. After work-up and purification by column chromatography, ethyl ester **59** was afforded in a 40% yield. This yield was lower than the one reported by Koskinen *et al.*[126] (i.e. 96%) but deemed to be sufficient for our purpose. The structure of the product was confirmed by ^{1}H and ^{13}C NMR spectral analysis. The appearance of a singlet at 3.71 ppm in the ^{1}H NMR spectrum was attributed to the new CH_2 attached to the oxygen atom of the alcohol. The presence of this new CH_2 group was further confirmed by ^{13}C NMR, with the appearance of a signal at 68.5 ppm.



Scheme 9

Finally, ester **59** was hydrolysed to the corresponding carboxylic acid **45** by treatment with 2 M aqueous sodium hydroxide in methanol at room temperature. After work-up and column chromatography purification, product **45** was obtained in 55% yield. The formation of carboxylic acid **45** was verified by ^{1}H NMR spectroscopy as the signals at 1.16 ppm and 4.10 ppm, corresponding to the ethyl ester group, were no longer visible. In addition, the broad singlet that appeared at 8.60 ppm was attributed to the proton signal of the carboxylic acid group.

At this point in the synthesis, the dibenzyl protecting group of **45** needed to be cleaved. Koskinen *et al.*[126] had carried out this reaction using H_2(g) at 3 atm. in the presence of a palladium on charcoal catalyst. As mentioned earlier, at this time the Chemistry department at Heriot-Watt University did not have a functioning Parr apparatus to enable us to perform similar reaction. Therefore, we attempted to remove this group at atmospheric pressure. Thus, carboxylic acid **45** was dissolved in absolute ethanol and to this was carefully added a small portion of 5% Pd/C catalyst. The black heterogeneous mixture was evacuated and purged with H_2(g) several times before being left to react at room temperature for two days (*Scheme 10*). After this time, the reaction was worked up and the crude product obtained was analysed by both ^{1}H and ^{13}C NMR. However, both spectra still showed the presence of two benzyl groups in the exact ratio as found for the starting material **45**. This implied that no reaction had occurred; we therefore sought alternative strategies for removal of benzyl groups.



Scheme 10

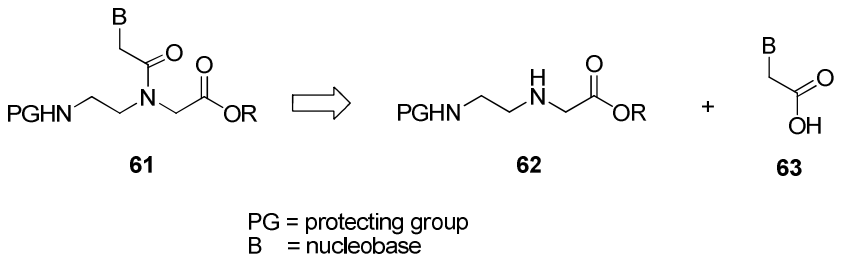
Catalytic transfer-hydrogenation (CTH) offers the possibility for an efficient cleavage of benzyl groups. CTH has successfully been applied to the removal of benzyl groups from amino functions[131] as well as hydroxyl moieties.[132] It has also found applications in the reduction of multiple bonds (alkenes, alkynes), carbonyl compounds, nitriles, imines, azo compounds, nitro compounds etc.[133] The reaction is accomplished with the aid of an organic molecule which acts as the hydrogen donor in the presence of a catalyst. Although this approach has been principally reported for use of palladium-based catalysts, other catalysts based on rhodium , nickel and platinum can also be used.[133] Using this

approach, there are many different hydrogen donors that have been employed for the removal of benzyl groups from amino functions. The most common are formic acid,[134] 1,4-cyclohexadiene,[135] hydrazine hydrate[136] and ammonium formate.[131, 137, 138]
Based on these reports, we reasoned that the methods described by Ram *et al.*[131, 138] could be utilised to remove the dibenzyl protecting group of **45**. Therefore, carboxylic acid **45** was dissolved in methanol and to this solution were carefully added one equivalent of 10% Pd/C and two equivalents of ammonium formate. The resulting heterogeneous mixture was heated at a gentle reflux and the progress of the reaction was carefully followed by TLC. Unfortunately, after five hours, TLC analysis still showed the presence of a UV active compound. Subsequent ^{1}H NMR spectral analysis of the crude product after work-up revealed a mixture of both the di-benzyl protected starting material **45** and the mono-benzylated intermediate **60** (*Scheme 10*). In an attempt to drive the reaction further, the CTH reaction was repeated on this crude mixture. However, after another five hours, the reaction failed to reach completion. Furthermore, none of the desired, completely de-benzylated, product **39** could be observed. Therefore, it appeared that CTH could not be utilised for this reaction.

The preparation of the Boc-protected hydrophilic spacer **40** was brought with difficulties in our hands. Although we were able to repeat the work described by Koskinen *et al.*,[126] the yields for the intermediates were not as high as reported. Due to instrumental limitations, we were not able to complete this reaction scheme and further investigations would be necessary to finish the reaction pathway. However, with time pressing on, we decided to take the limited amounts of the Boc-protected intermediate **47** (*Scheme 5*) and the dibenzylated intermediate **45** (*Scheme 9*) obtained from this work and use them in the next stage of this project, i.e. the preparation of lipid-spacer-PNA monomer models. We reasoned that if these studies proved fruitful, we could revisit the preparation of the Boc-protected AEEA spacer in a later stage.

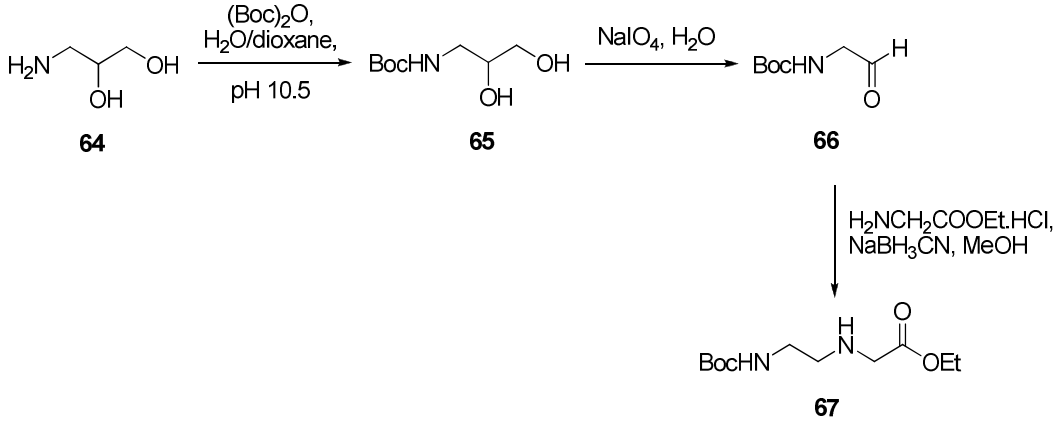
2.2.2 Preparation of PNA monomers

The preparation of PNA monomers has been extensively reported in the literature since their development by Nielsen *et al.*[40] in 1991. The commonest synthetic pathway used to prepare the classical PNA monomers, such as **61,** is highlighted in *Scheme 11*. The key step involves linking the protected *N*-(2-amino)glycine ester backbone (**59**) to the acetic acid derivative of the appropriate nucleobase (**60**).



Scheme 11

During the course of the earlier studies conducted by Murray,[2] a number of the strategies reported in the literature for the preparation of the PNA backbone **62** were evaluated. From this work, the approach developed by Finn *et al.*,[139] leading to the Boc-protected PNA backbone, gave the best results. This strategy starts with the Boc protection of commercially available 3-amino-1,2-propanediol **64** using di-*tert*-butyl dicarbonate, affording diol **65** in quantitative yield. Periodate cleavage of the vicinal diol **65** leads to the Boc protected amino aldehyde **66** in high yield, reductive amination of which with cyanoborohydride and ethyl glycinate affords *N*-(2-Boc-amino-ethyl)glycinate **67** (*Scheme 12*).[139]

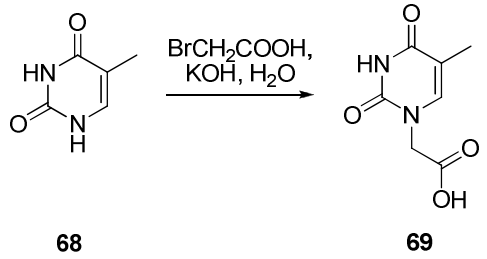


Scheme 12

This strategy was therefore repeated in our case and comparable yield to Finn *et al.* were obtained for backbone **67** (i.e. 55%).

Now that the PNA monomer backbone had been successfully prepared, we needed to prepare the nucleobase acetic acid derivatives next. At this stage in the project, we decided to concentrate our efforts on the preparation of the thyminyl PNA monomer because this nucleobase does not have any exocyclic nucleophilic functionality that could interfere in the later steps of the synthetic pathways.

The preparation of thyminyl acetic acid **69** is outlined in *Scheme 13*. This involved treatment of an aqueous solution of commercially available thymine (**68**) with bromoacetic acid in the presence of potassium hydroxide, as reported by Kosynkina *et al.*.[140] This methodology had also been used by Murray.[2] This reaction was successfully reproduced here to yield thyminyl acetic acid **69** in 97% yield, following filtration of the acidic aqueous reaction mixture.



68 **69**

Scheme 13

At this point, we were ready to connect the thyminyl acetic acid derivative **69** to the Boc-protected PNA backbone **67**. In order to carry out this reaction, the carboxylic acid function of derivative **69** required first to be activated. A range of different strategies have been reported in the literature to achieve this activation. *In situ* activation using 3-hydroxy-1,2,3-benzotriazin-4(*3H*)-one (DhbtOH) and *N,N'*-dicyclohexylcarbodiimide (DCC) was one of the first methods published (*Figure 50*). This was found to be very efficient[1, 141] for the introduction of thyminyl, cytosinyl and adeninyl derivatives. In another work, the guaninyl derivative was reported to be efficiently attached to the PNA backbone *via* condensation using bromotris(pyrrolidino)phosphonium hexafluorophosphate (PyBroP®) as the activating agent.[141] More recently, the use of uronium and aminium salts such as HATU, HBTU and TBTU, have been reported to give higher yields that the originally DhbtOH/DCC strategy.[141-144] TBTU[145] and HBTU[146] are two most

popular in situ activation reagents used in solid phase and solution phase peptide synthesis.[147]

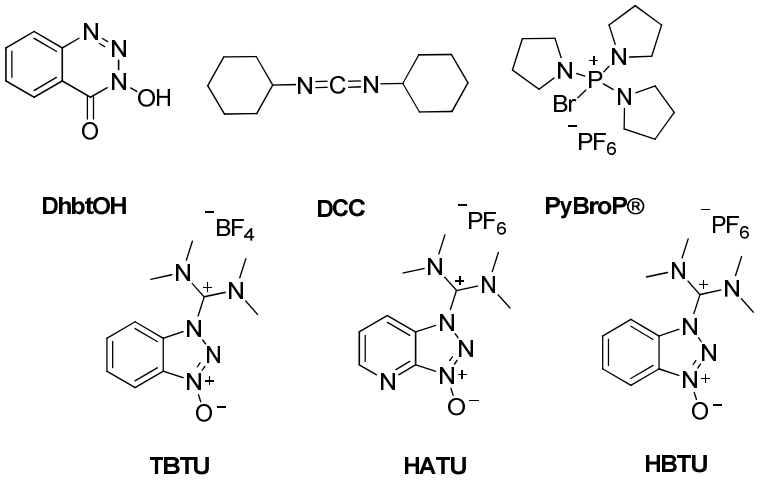
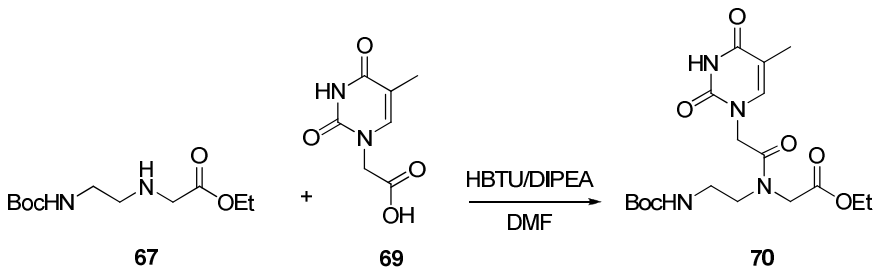


Figure 50

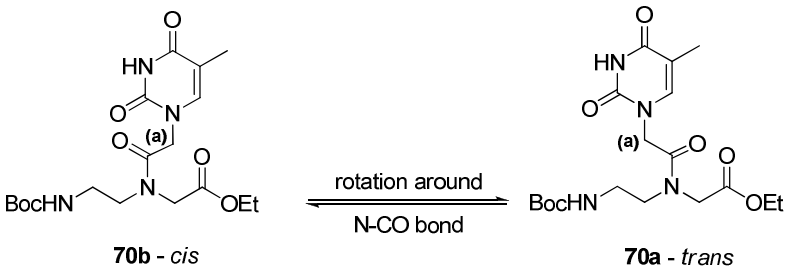
For our purpose, we decided to follow the work of Murray[2] and chose to couple thyminyl acetic acid **69** to the PNA backbone **67** using HBTU as the activating agent. After work-up and purification by column chromatography, the desired thyminyl PNA monomer **70** was obtained in 70% yield (*Scheme 14*).



Scheme 14

Finn and co-workers[139] have reported high barriers of rotation (10-25 kcal·M^{-1})[148] and low exchange rate (0.5-2 s^{-1}, 37 ^{o}C)[148] around the tertiary amide bond of **70**, resulting in a two rotational isomers, *cis*:*trans*, in a 1:2 ratio (*Scheme 14*). ^{1}H NMR studies carried out by Chen[149] and Brown[150] (separately) have shown that the major form is the *trans* rotational isomer (**70a**), where the methylene protons (**a**) next to the nucleobase are toward the 2-aminoethyl unit. In the minor *cis* isomer (**70b**), they are on the side of the glycyl unit. This result had been obtained using 2D NOESY[149] spectra, showing the in-

teractions between the protons of the methylene protons **(a)** (*Scheme 15*) with the methylene protons present on the PNA backbone.



70b - *cis* **70a** - *trans*

Scheme 15

We, likewise, observed the presence of two rotamers in the ^{1}H and ^{13}C NMR spectra recorded for **70**. Their presence was clearly indicated by the "splitting" of certain NMR signals, as exemplified in the ^{1}H NMR spectrum shown in *Figure 51*. For example, the CH_2 **a** (*Scheme 15*) was split into two singlets; a smaller one at 4.45 ppm and a larger one at 4.60 ppm. The ratio of the major rotamer to the minor rotamer was found to be 2:1, which was consistent with the report of Finn *et al.*.[139]

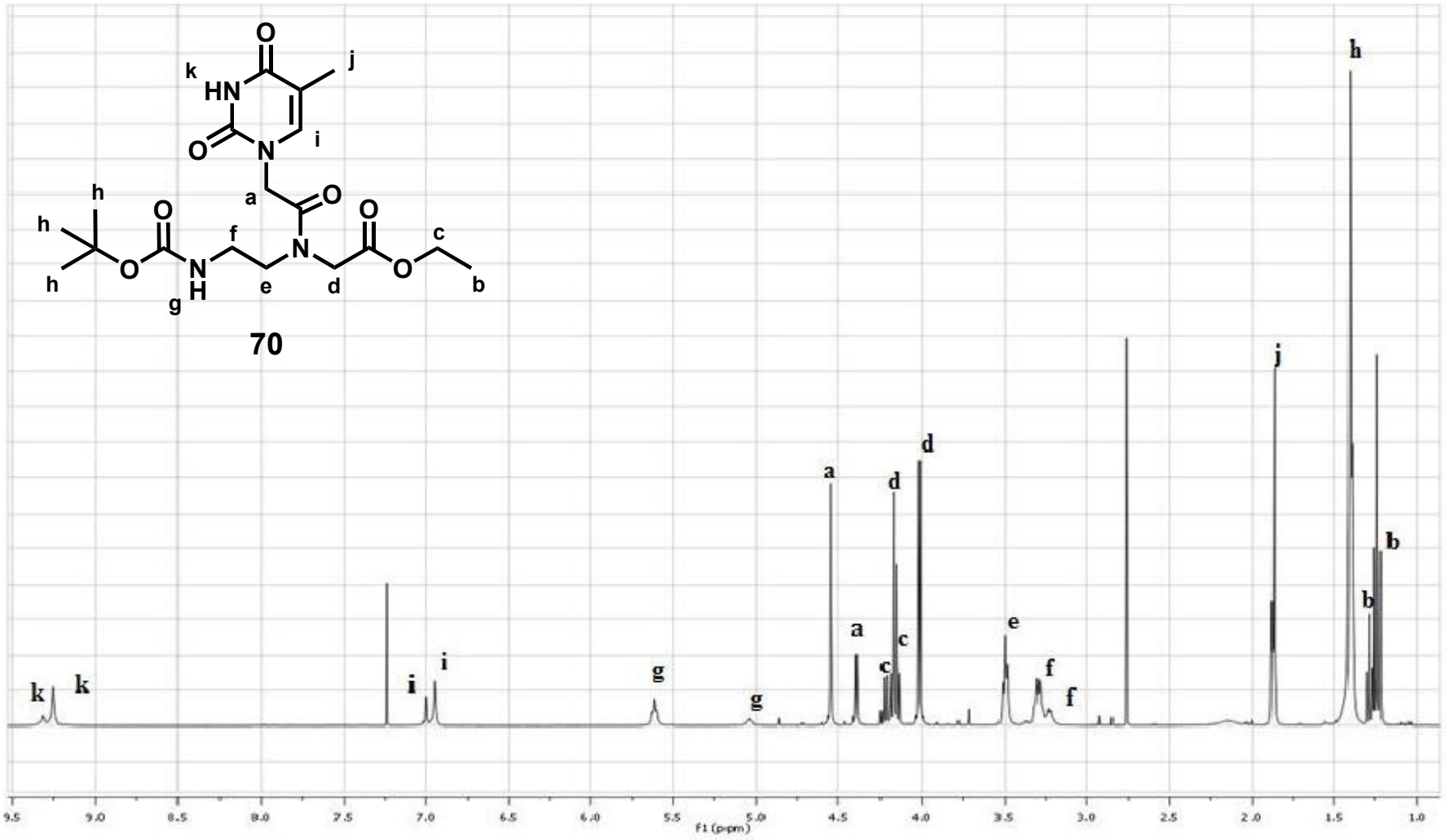


*Figure 51 – ^{1}H NMR of **70** (400 MHz, $CDCl_3$)*

The preparation of thyminyl PNA monomer **70** was further confirmed using X-ray crystallography. A suitable crystal of **70** was grown for these studies from diethyl ether and methanol using the vapour diffusion method at room temperature. Compound **70** crystallised in the monoclinic space group, P2(1)/n, with one molecule in the asymmetric unit as shown in *Figure 52*. The full crystallographic data are shown in *Appendix A*. Although in solution the tertiary amide bond in **70** exists as a rotameric mixture, in the crystal structure only the *trans* rotamers was observed (i.e. where the carbonyl was directed towards the glycyl component of the PNA backbone).

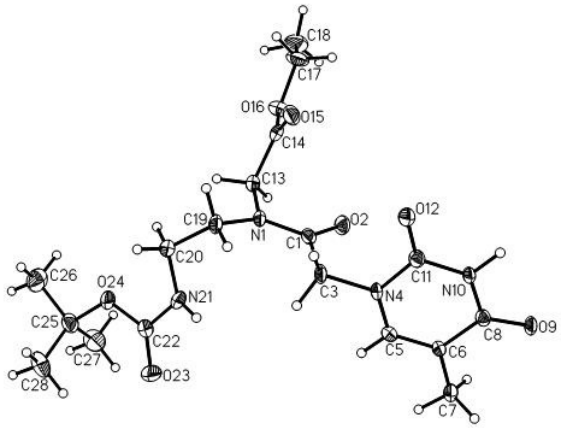


Figure 52 – ORTEP view of ***70***

The packing analysis, shown in *Figure 53*, revealed that **70** exhibited the anticipated centrosymmetric hydrogen-bonded dimeric structure with an H···O distance of 1.86 Å.

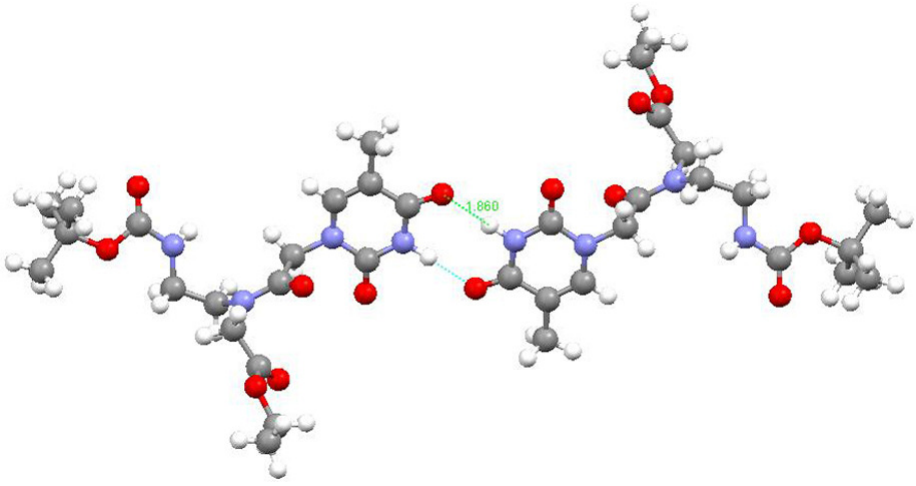


Figure 53

Such arrangement leads to π-stacking of the nucleobase pairs in a stepwise manner, as depicted in *Figure 54*. The vertical distance between the two carbonyl groups of two adjacent thymine rings in the stack was found to be 6.89 Å.

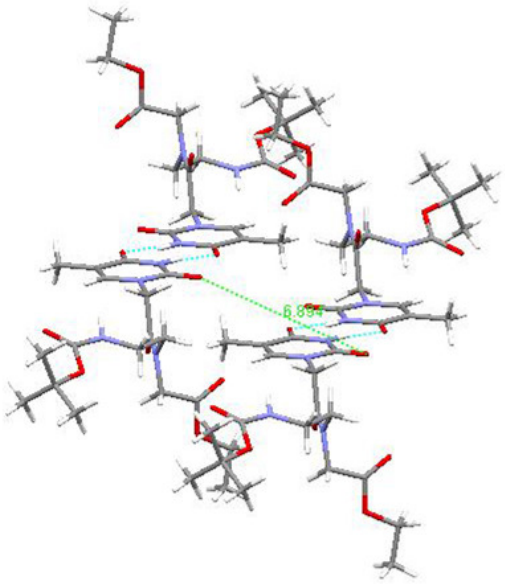


Figure 54

Finally, the adjacent molecules were held together by the formation of two N-H···O hydrogen bonds. One was formed between the NH of the PNA backbone and the oxygen atom of the neighbouring ester carbonyl group; the H···O distance was 2.83 Å. The second was formed between the same NH group and the C_{11} carbonyl oxygen of the neighbouring thyminyl ring ($H\cdots O_{12}$ distance = 2.36 Å - *Figure 55*).

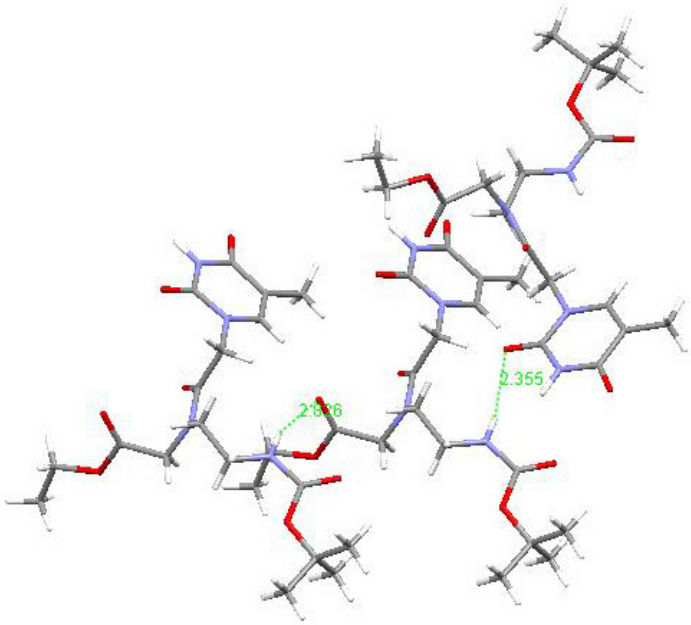
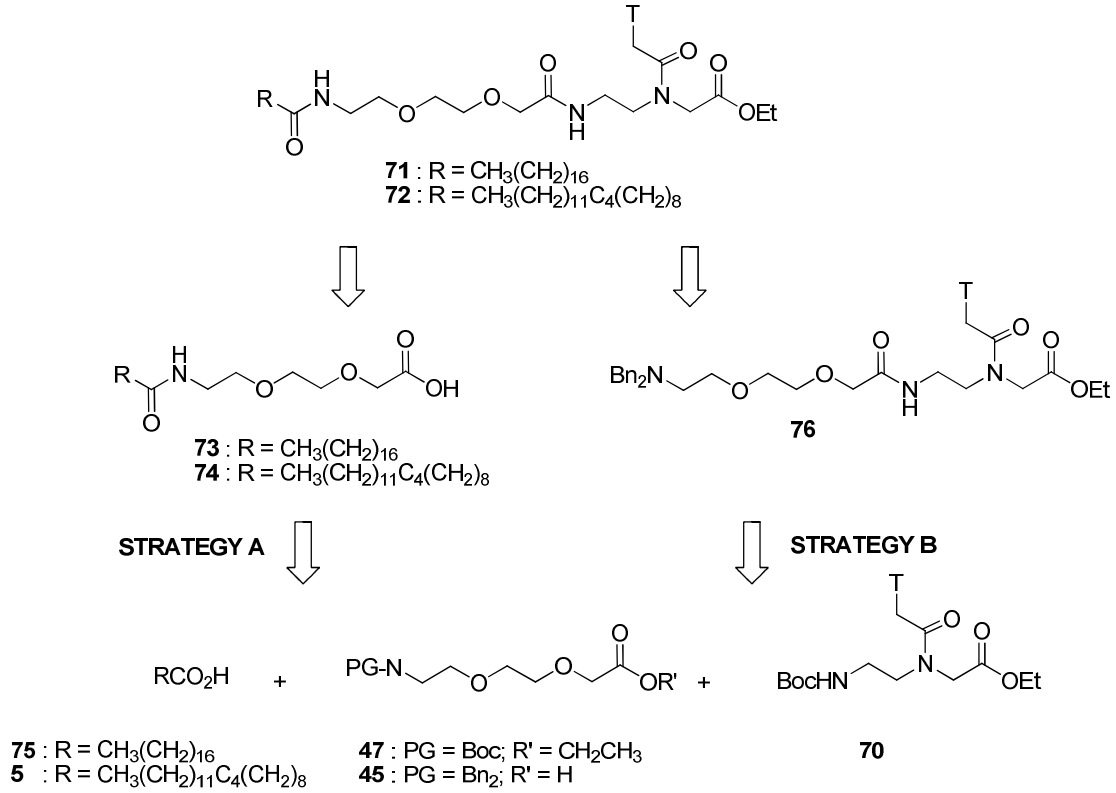


Figure 55

2.2.3 Preparation of lipid-'PEG-like' spacer-PNA monomer models

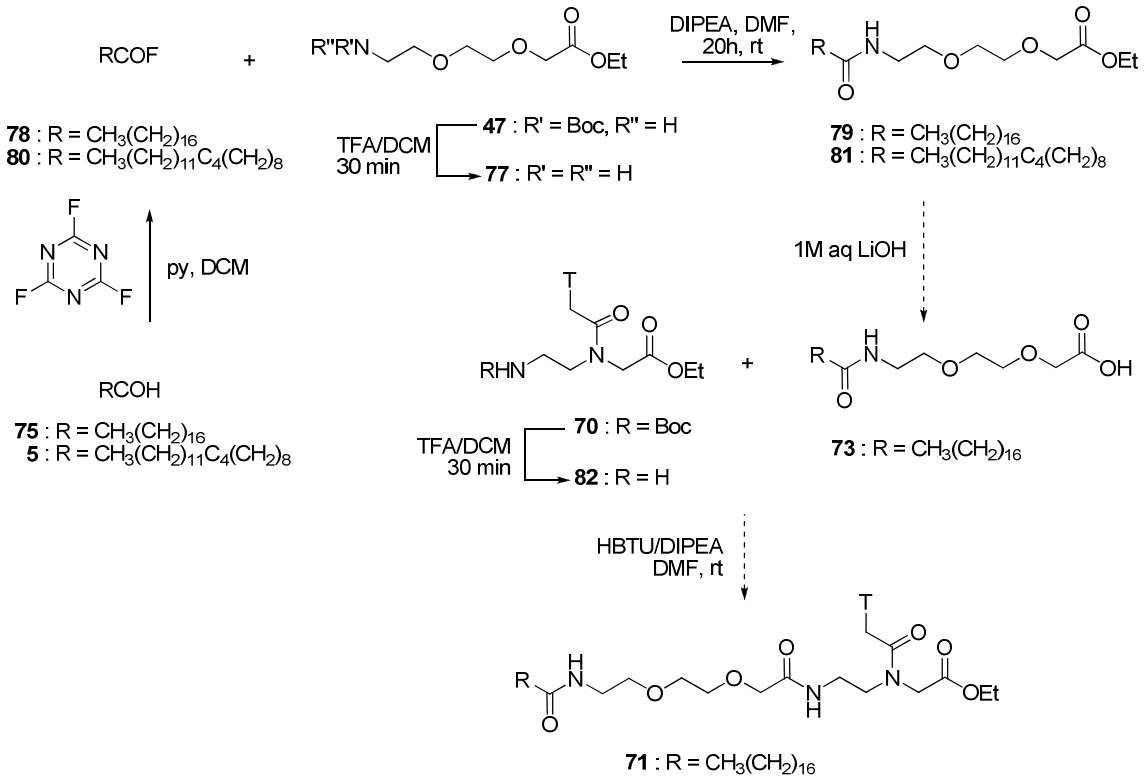
The retrosynthetic analysis outlined in *Scheme 16* identified two potential strategies for the preparation of lipid-'PEG-like' spacer-PNA monomer models **71** and **72** from the three components, i.e. the thyminyl PNA monomer **70**, the AEEA spacer intermediates **45** and **47**, and the lipids **5** and **75**. We proposed to covalently attach the PNA monomer to the carboxylic acid terminus of the AEEA spacer and the lipid to the amino terminus through the formation of amide bonds.



Scheme 16

As highlighted in Section 2.2.1, the synthesis of Boc AEEA spacer **40** (*Scheme 5*) had proved more difficult than expected and the synthetic route had not yet been completely devised. However, during the course of our research to develop a viable pathway to **40**, we had managed to prepare a small amount of **47**. This was sufficient to enable us to start exploring the use of this spacer in the construction of the desired lipid-spacer-PNA conjugates.

Our initial investigations centred on Strategy A and the first step involved removal of the Boc protecting group from **47** (*Scheme 17*).



Scheme 17

The removal of the Boc group from **47** was successfully achieved using a mixture of TFA and dichloromethane (1:1 v/v) to afford the corresponding trifluoroacetic acid salt **77** in a quantitative yield. The next step in Strategy A involved the attachment of the lipid chain to the AEEA spacer. Before being able to accomplish this, the carboxylic acid of the lipid needed to be activated. Murray[2] had prepared the acid fluoride derivative of commercially available stearic acid (**75**) and 10,12-pentacosa-diynoic acid (**5**) to attach the lipids to his PNA monomers. Therefore, we decided to adopt the same methodology here. A solution of **75** in DCM was treated with cyanuric fluoride in the presence of pyridine and the resulting mixture was heated to gentle reflux for three hours. The desired acid fluoride **78** was obtained as a waxy solid in 92% yield after an aqueous work-up. The ^{1}H NMR spectrum showed the expected peaks for **78** with the signal assigned to the CH_2 group adjacent to the carbonyl function, moving downfield from 2.35 ppm in the spectrum recorded for **75** to 2.49 ppm. The multiplicity of this signal also increased to a doublet of triplets due to the methylene protons coupling to the fluorine atom as well as the neighbouring CH_2 protons. The ^{13}C NMR spectrum also showed the

expected peaks; the carbonyl carbon of **78** appeared as a doublet at 163.8 ppm due to coupling to the neighbouring fluorine atom. Additionally, in the IR spectrum of **75**, we observed that the absorbance for the carbonyl group had moved from 1681 cm^{-1} for the acid to 1846 cm^{-1} which is characteristic of an acid fluoride function. These results were comparable with the report by Murray.[2]

Having successfully prepared steraoyl fluoride **78**, this was now ready to be coupled to to **77** (*Scheme 17*). Thus, a solution **77** in DMF was treated with DIPEA and then added to a solution of **78** in dichloromethane. The resulting mixture was allowed to stir overnight at room temperature. After work-up and purification by column chromatography, the lipid-spacer adduct **79** was obtained in 79% yield. ^{1}H and ^{13}C NMR spectra recorded for **79** showed all the expected signals apart for the ethyl ester CH_3 that was hidden by the large signal corresponding to the aliphatic CH_2 of the lipid chain. The appearance of the broad triplet at 5.95 ppm on the ^{1}H NMR spectrum was assigned to the NH proton of the new peptide bond.

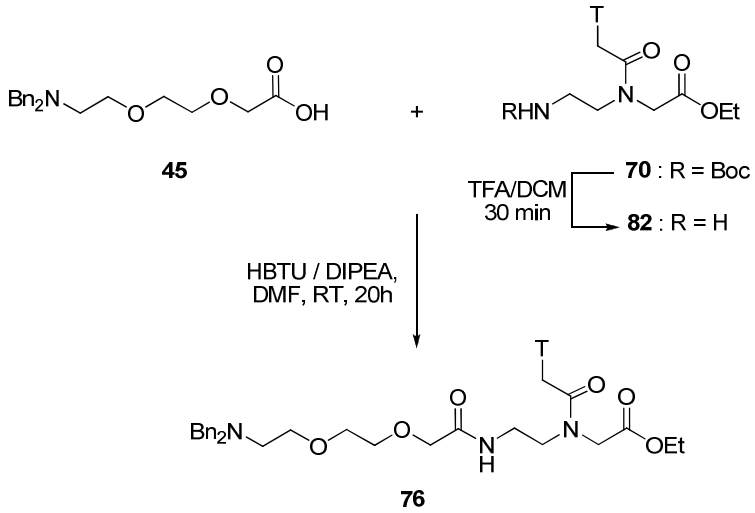
Strategy A was also followed for the attachment of the spacer to 10,12-PCDA (**5**). Again the commercially available lipid **5** was first converted into its acid fluoride derivative **80** following the same procedure described for the preparation of **78**. After work-up, **80** was obtained in 82% yield. **80** was then successfully coupled to **77** using a similar protocol to that employed for the preparation of saturated analogue **79**. After work-up and purification by column chromatography, **81** was obtained in 69% yield. Again, the formation of **81** was confirmed by ^{1}H and ^{13}C NMR spectroscopy. As it was the case for **79**, the appearance of the broad triplet at 6.29 ppm on the ^{1}H NMR spectrum corresponding to the NH proton of the new peptide bond. Unfortunately, diacetylene **81** was found to be unstable in its solid state; it quickly turned from colourless to blue, as a result of topochemical polymerisation, even in the dark at -20 ^{o}C. This precluded any further chemical manipulations of **81**. It therefore appeared that Strategy A was not applicable for the preparation of diacetylenic lipid-PEG-like spacer-PNA monomer model **72** (*Scheme 16*).

The next planned steps in Strategy A involved hydrolysis of the ethyl ester of **79** to yield the corresponding carboxylic acid **73**. This would then be coupled to the free *N*-terminus PNA monomer **82** using HBTU as the activating reagent to afford model **71**. Unfortunately, due to the small amounts of **47** collected and used for the preparation of

77 only small amounts of **79** and **81** were obtained. This was deemed insufficient for the final two steps required for the preparation of models **68** and **69** and these were therefore not attempted.

We have also carried out some preliminary investigations into the utility of Strategy B for the preparation of the lipid-spacer-PNA monomer models **71** and **72**. For these studies we chose to use the dibenzyl protected AEEA spacer **45** (*Scheme 16*) because we had a small quantity of this compound in the lab. This model reaction was only used to identify which strategy would be more applicable for the preparation of models **71** and **72**. We reasoned that, once we had developed a more viable route to Boc-protected spacer **40**, it could be easily substituted for **45**.

The first step in this model strategy involved the removal of the Boc protecting group from the *N*-terminus of PNA monomer **70**, under acidic conditions, to afford the trifluoroacetic acid salt derivative **82** in a quantitative yield (*Scheme 18*). In order to couple **82** to carboxylic acid **45**, we decided to use HBTU as the activating agent in the presence of DIPEA. However, after work-up, a complex crude mixture was obtained. Careful purification of this crude product by flash column chromatography, finally afforded **76** in the very low yield of 6%. The reason for this low yield is unknown and due to the lack of any further amount of **45**, we were unable to repeat the reaction.

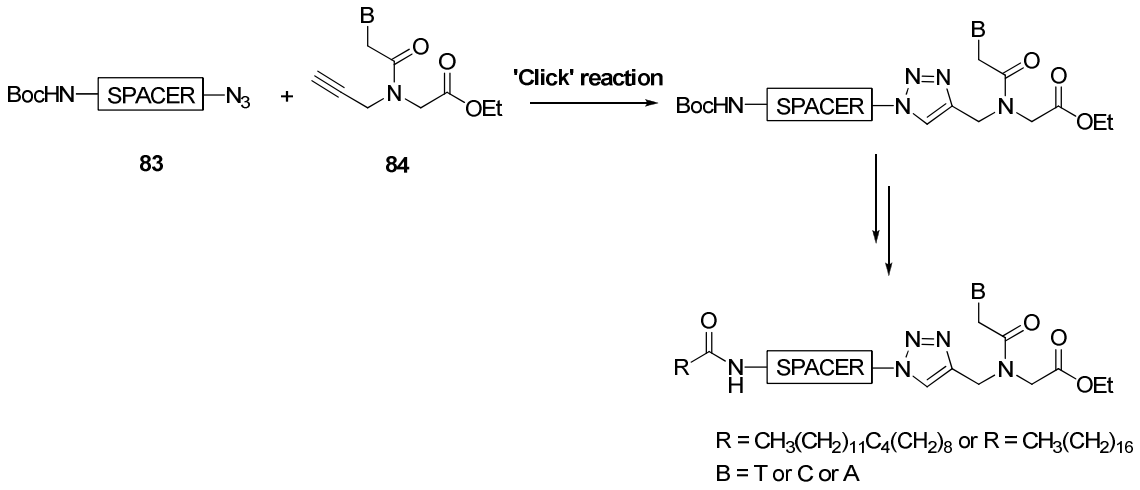


Scheme 18

Thus it appeared from our Strategy A model studies, that the PEG-like spacer could be coupled to the lipid tail through amide bond formation in reasonable yield. However,

from our Strategy B model studies, it was concluded that peptide coupling was not adapted to connect the PEG-like spacer to the PNA monomer due to the low yield obtained.

At this stage of the project, it became apparent that we needed to re-evaluate our initial strategies. The problems of developing a viable synthetic route to the required Boc-AEEA spacer **40**, coupled with the low yield obtained when trying to attach the PNA monomer to the spacer through amide bond formation meant that our original plans were unworkable. We therefore started to look for other approaches to couple the spacer to the PNA monomer which would not involve peptide bond formation. We reasoned that the new approach sought should avoid the requirement for the synthesis of the Boc-PEG-CO_2H spacer **40** (*Scheme 5*). We became aware of literature reports based on the 'Click' reaction. This novel reaction is a copper(I)-catalysed version of the Huisgen 1,3-dipolar [3,2] cycloaddition involving terminal alkynes and organic azides. The corresponding 1,4-disubstituted [1,2,3]-triazoles-containing products are formed both chemo- and regioselectively in high yields. These triazoles have generated large biological interest as they have been found to mimic amide bonds. The preparation of these triazole linkages will be discussed in the next section (Section 2.3). Bearing this in mind, we therefore decided to embark upon the investigation into the utility of this reaction for coupling of the PEG-like spacer to the *N*-terminus of PNA monomers (*Scheme 19*).



Scheme 19

As shown in *Scheme 19*, it was decided that the spacer would contain the azido moiety while the PNA monomer would be functionalised at its *N*-terminus with an alkynyl function. Therefore, synthetic pathways to both the azido component **83** and the *N*-

alkynyl PNA monomer **84** have to be developed. This time, we decided to concentrate on preparing model conjugates containing thyminyl, adeninyl and cytosinyl PNA monomers to validate our proposal. If successful, PNA oligomers would be incorporated at a later stage. The development towards these goals will be discussed in detail in Section 2.4 of this chapter.

2.3 DI-SUBSTITUTED 1,2,3-TRIAZOLE

2.3.1 Background

The [1,2,3]-triazole moiety can be found in biologically active molecules including *anti*-HIV activity,[151] *anti*-microbial activity against Gram positive bacteria,[152] potent *anti*-histaminic,[153] *anti*-platelet agents[154] and tuberculosis inhibition activity[155] (*Figure 56*). This chemical function is also widely found in industrial materials such as dyes and agrochemicals.

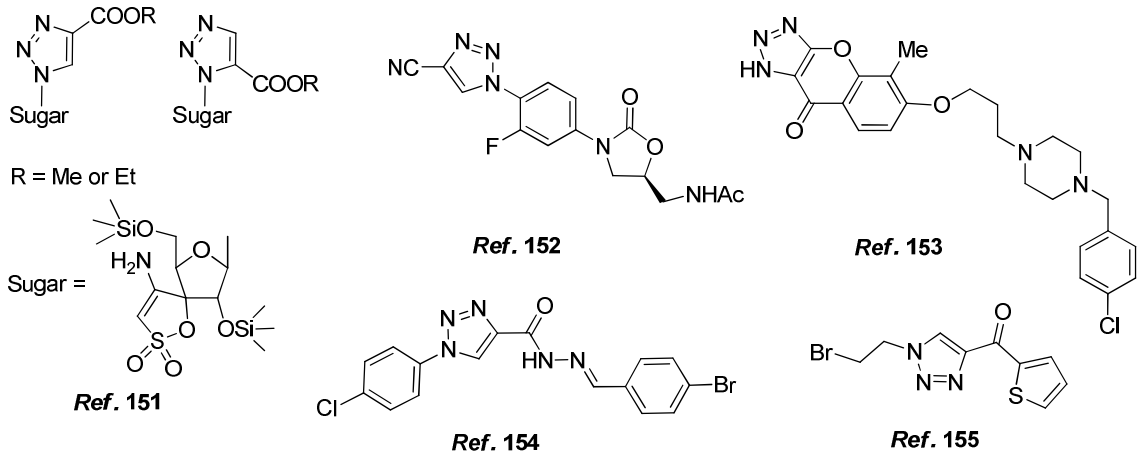
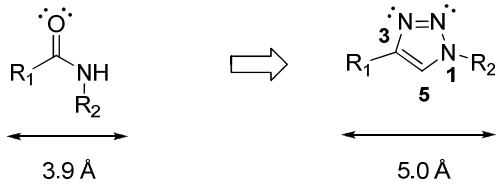


Figure 56 – examples of some biologically active molecules containing [1,2,3]-triazoles

[1,2,3]-Triazole occurrence may be explained by the fact that these appear to mimic peptide bonds, i.e. they are a *cis* peptide bioisostere. This rigid heterocyclic function has been shown to imitate the atom placement and electronic properties of a peptide bond without holding the same susceptibility to cleavage. [1,2,3]-triazoles are also nearly impossible to oxidise or reduce.[156]

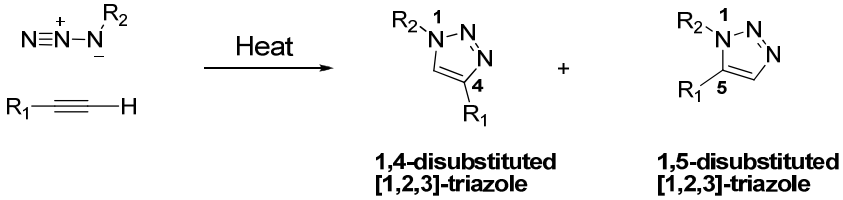
It is rather surprising that [1,2,3]-triazoles should act as peptide mimics as, as shown in *Scheme 20*, the triazole backbone contains an additional atom which increases the dis-

tance between R_1 and R_2 by 1.1 Å. However, van Maarseveen *et al.*[156] have stated that the peptide bond mimicry of [1,2,3]-triazoles is facilitated by the fact that they possess a much stronger dipole moment than an amide bond. These researchers reported that both the N(2) and N(3) triazolyl atoms could act as hydrogen-bond acceptors while the strong dipole could polarize the C(5) proton so that it functions as a hydrogen-bond donor, similar to the N-H proton of amides.[156]



Scheme 20

The commonest route for the synthesis of [1,2,3]-triazoles involves the well known 1,3-dipolar cycloaddition between azides and terminal alkynes as established by Huisgen in 1963 (*Scheme 21*).[157]



Scheme 21

Azides and alkynes are relatively stable under a variety of conditions and they are inert towards most biological and organic conditions, particularly water and oxygen. Thanks to their stability, azides remain invisible unless a good dipolarophile (such as an alkyne) is presented. However, elevated temperatures and long reaction times are often required and the result of this 1,3-dipolar cycloaddition is usually a mixture of the 1,4 and 1,5 regioisomers of [1,2,3]-triazole in approximately 1:1 ratio (*Scheme 21).*

In addition to the regioselectivity issue, another problem with this reaction is that azides can decompose exothermically, with loss of N_2, and therefore must be treated with care. Although aliphatic azides have a particularly high kinetic stability relative to other azides, the presence of other energetic functional groups in the molecule, such as alkynes, increases this hazard. Probably due to this safety hazard, the Huisgen 1,3-dipolar

cycloaddition between azides and alkynes had, until recently, never been exploited in depth.

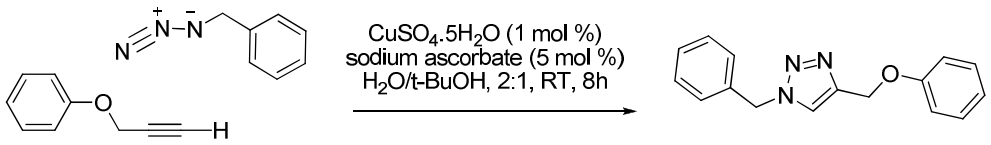
2.3.2 The copper(I) catalyzed "click" reaction

Renewed interest in the 1,3-dipolar cycloaddition, described above involving azides and organic alkynes, has occurred in the last seven years due to the findings reported separately by Sharpless[158] and Meldal.[159] In 2002, these researchers independently reported the use of a copper(I) catalyst in this reaction to improve product yield, decrease the time of reaction and eliminate the need for elevated temperatures. Furthermore, they discovered that this catalysed variant of 1,3-dipolar addition improved the 1,4- *versus* 1,5-regioselectivity. Due to the key features exhibited by the reaction, this copper(I)-catalysed [2+3] cycloaddition has been classified within the family of "click" reactions and it has become one of the most popular reactions for preparing [1,2,3]-triazole functions. Today, "click chemistry" as become synonymous with this specific copper(I)-catalysed [2+3] cycloaddition.

Click chemistry processes are a family of chemical reactions that were first described by Sharpless *et al.*[160] in 2001. These reactions were defined as being "*modular, wide in scope, giving high yields, generating only inoffensive by-products, stereospecific and simple to perform*".[160] In order to be considered as a "click" reaction, the process has to be insensitive to oxygen and water, use readily available starting materials and reagents, employ easily removed or non-hazardous solvents, i.e water and finally the product isolation must be simple, i.e. non chromatographic. Many types of reactions that adhere to these principles of "click" chemistry have since been described, including cycloadditions involving unsaturated species (e.g. hetero-Diels-Alder and 1,3-dipolar cycloadditions), nucleophilic substitutions (e.g. ring opening of epoxides, aziridines, aziridinium ions and episulfonium ions), "non-aldol" type carbonyl reactions (e.g. formation of ureas, thioureas, aromatic heterocycles, oxime ethers, hydrazones and amide) and additions to carbon-carbon multiple bonds (e.g. epoxidation, dehydroxylation, aziridination and sulfenyl halide addition, as well as Michael additions of Nu-H reactants).[160]

In the case of the Huisgen-1,3-dipolar "click" variant, numerous sources of copper(I) catalysts have been investigated. Two different experimental strategies have been described by Sharpless *et al.*.[161] In the first one, the active copper(I) catalyst was produced

in situ from a copper(II) salt (i.e. $CuSO_4 \cdot 5\ H_2O$) *via* reduction with ascorbic acid or sodium ascorbate in excess, as outlined in *Scheme 22*.[158] The presence of excess sodium ascorbate was found to prevent the formation of oxidative coupling side products that have often been observed when a copper(I) source (i.e. CuI) was directly used.[161]



Scheme 22

In the second procedure, the catalyst was introduced in the form of the copper wire or shavings. The active Cu(I) catalyst was formed *via* comproportionation of the Cu(II)/Cu(0) couple.[161] This method has proved to be convenient in biosynthesis where substrates do not tolerate sodium ascorbate or its oxidation products. [162]

As mentioned above, in the presence of a copper(I) catalyst the 1,3-dipolar cycloaddition of a variety of terminal alkynes and organic azides has resulted in the regioselective formation of 1,4-disubstituted-[1,2,3]-triazoles. The reaction has been shown to be tolerant to a wide range of pH (*ca.* 4 to 12), resistant to a broad range of temperatures (from 0 to 160 ^{o}C), and finally, to proceed in a variety of solvent systems (with a preference for aqueous systems, including serum and whole blood[161]). Moreover, the other functional groups present on the reactant do not require protection as they do not interfere with the reaction. The corresponding triazole cycloadducts can usually be isolated in a pure form in high yields by simple filtration or extraction. Since the foundation of this type of "click" reaction have been laid, there has been an exponential growth in publications in the literature describing its applications.[163] Two examples that have been recently reported are described below due to their close relation to the work being undertaken in this research project.

Park and Lee[164] have applied the "click" reaction to verify the successful immobilisation of PNA oligomers onto microarrays. This involved first derivatising an amine-activated glass slide by incubation with [(4-ethynylphenylcarbamoyl)-methoxy]-acetic acid (4EPA) in DMA containing HBTU and DIPEA (*Figure 57*).[164] The immobilisation of the PNA was then achieved by treating the phenylacetylene-derivatised glass slide with (3-azido-2-oxo-2H-chromen-7-yloxy)-acetic acid (AZCO)-linked PNA oligomers, un-

der standard "click" reaction conditions. The spots, thus, obtained were highly fluorescent due to the presence of conjugated aromatic groups. Therefore, the successful immobilization was verified through the non-destructive and direct fluorescence method.[164] They therefore showed the possibility to couple PNA oligomers to solid substrates using the highly efficient "click" reaction.

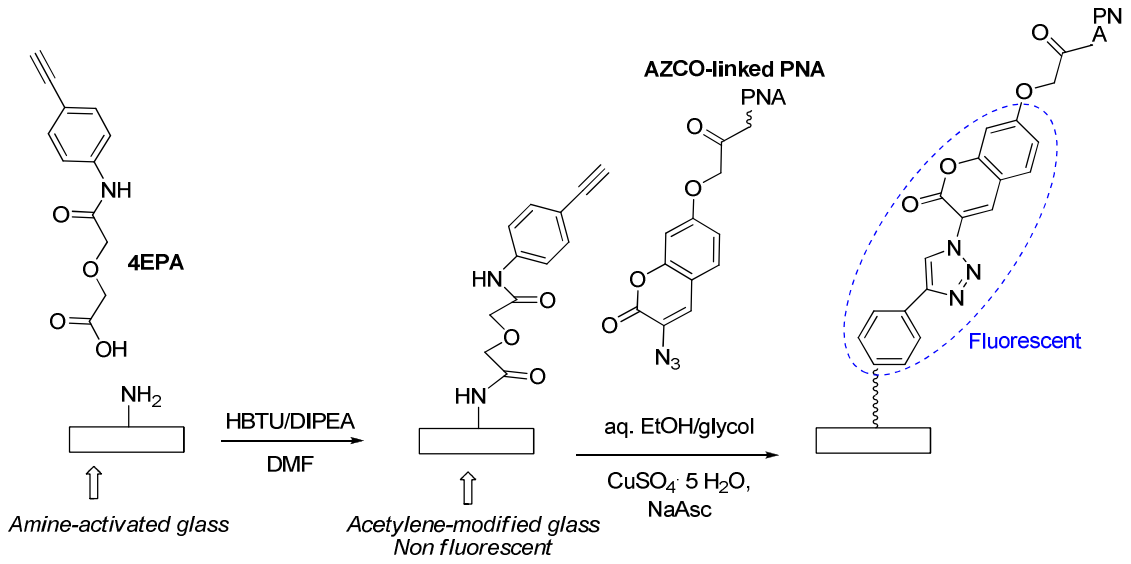
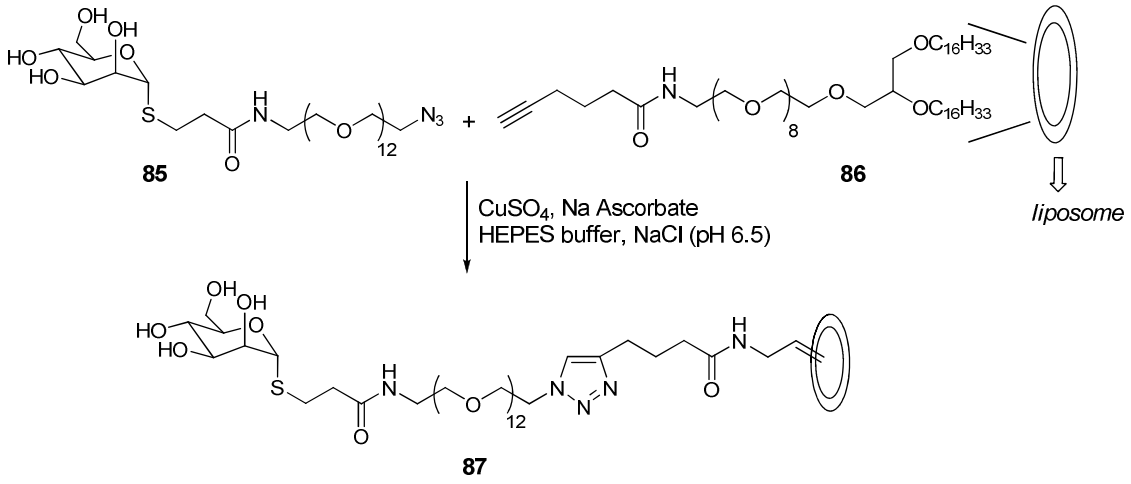


Figure 57

'Click' chemistry has also been used by Schuber *et al.*[165] for the preparation of mannose-functionalised liposome **87** (*Scheme 23*).



Scheme 23

In their work, an unprotected α-D-mannosyl derivative **85** carrying a spacer functionalised with an azide group was reacted at the surface of preformed liposomes **86** which incorporated a synthetic lipid bearing a terminal triple bond.[165] The integrity of the

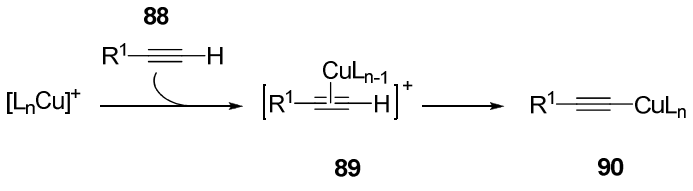
newly functionalised liposomes **87** was then verified using a dynamic light scattering technique. The research group also showed that the reaction did not affect the mannose now present at the surface of the vesicles and that these residues could be engaged into further interactions.[165]

These two examples were relevant for our work, as we aimed to construct liposomes from [1,2,3]-triazoles-containing lipids and we required the head-groups present on the outer surface to be readily accessible for later investigations.

2.3.3 Mechanism of copper(I)-catalysed Huisgen cycloaddition

Computational studies have been carried out in order to try to identify the roles of the copper(I) catalyst. It has been known for a long time that the uncatalysed Huisgen cycloaddition is a concerted process. The energy barriers for the formation of both 1,4- and 1,5-regioisomers have been found to be very close, 25.7 and 26.0 kcal / mol respectively. As mentioned earlier, the experimental data collected support these calculation, as *ca.* 1:1 mixture of the two regioisomers are typically obtained.[161] However, these studies have also shown that the copper(I) catalysed variant cannot occur in a concerted fashion.

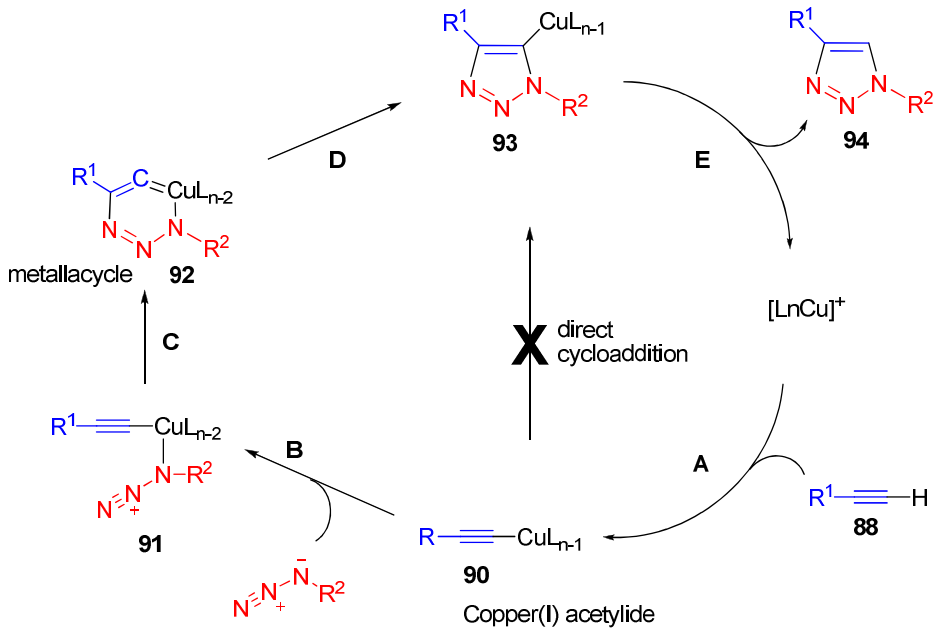
In light of this, Sharpless *et al.*[161] have proposed the mechanism outlined in *Scheme 25*. This suggestion has been supported by DFT calculations and experimental evidence. Thus, the sequence of this metal catalysed process begins with the coordination of the alkyne **88** to the copper(I) species (step **A** – *Scheme 24*). It has been reasoned that species **90** (*Scheme 24*) is formed *via* the initial π-complex formation **89**, in which one of the ligands L has been displaced by the incoming terminal alkyne.



Scheme 24

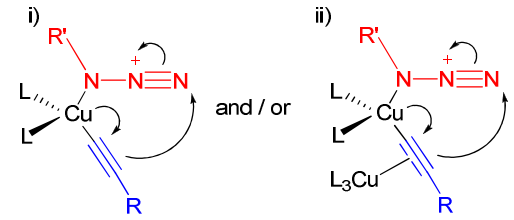
Conversion of the alkyne **88** to the acetylide **90** is known to be involved in many C-C bond forming reactions where Cu acetylide species **89** are intermediates. This initial co-

ordination to form **89** lowers the pK_a value of alkyne C-*H* by up to 9.8 units, making the attack onto the C-H bond possible.[161] Although the coordination of the acetylene to the copper(I) (without its deprotection) should activates a mechanism toward a 1,3-dipolar cycloaddition, Sharpless *et al.*[161] found that the calculated the barrier for this process to be at 27.8 kcal / mol. And the value for this new barrier is actually higher than the barrier for the uncatalysed process (*ca.* 26 kcal / mol). Among direct cycloaddition alternatives, the one with the neutral acetylide **90** resulted in an energy barrier calculated at 23.7 kcal / mol. Therefore, from these results, it was evident that the direct cycloaddition had to be ruled out.[161]

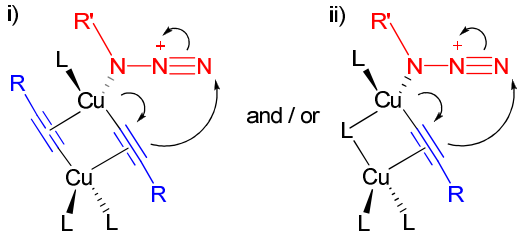


Scheme 25 – Mechanism proposed by Noodleman, Sharpless and Fokin [161]

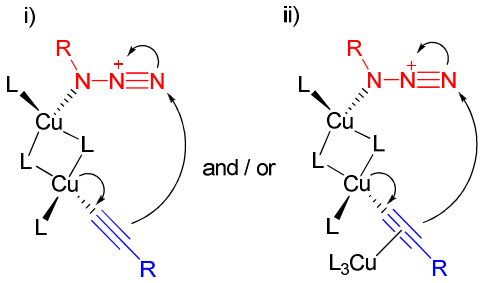
Thus, Sharpless *et al.* proposed that in the next step (**B** – *Scheme 25*) the azide replaces one of the ligands and binds to the copper atom *via* the nitrogen proximal to carbon, forming intermediate **91**. The exact nature of this reactive intermediate is still unknown but recently, Leigh *et al.*[166] have proposed various structures, based on recent DFT studies. These are summarized in *Figure 58.* This research group have proposed that it may be that the nature of the exact intermediate involved in the reaction pathways may depend upon the conditions employed, e.g. solvents, bulk and coordination number of the ligand added.



Type A: i) one copper atom or ii) one copper atom and π-activation of Cu-acetylide by a second copper atom.



Type B: i) Bis-alkyne bridged intermediate, or ii) two μ-coordinated copper atoms featuring a π-activated Cu-acetylide. These studies differ from A(ii) only by two copper atoms being double bridged. Doubly-bridged intermediates are not possible with bidentate ligands for copper (assuming RN_3 coordinates to Cu).



Type C: The azide and the alkyne are on different copper atoms. i) non-π-activated alkyne or ii) π-activated alkyne. These differ from B only in the location of the alkyne which reacts with the azide.

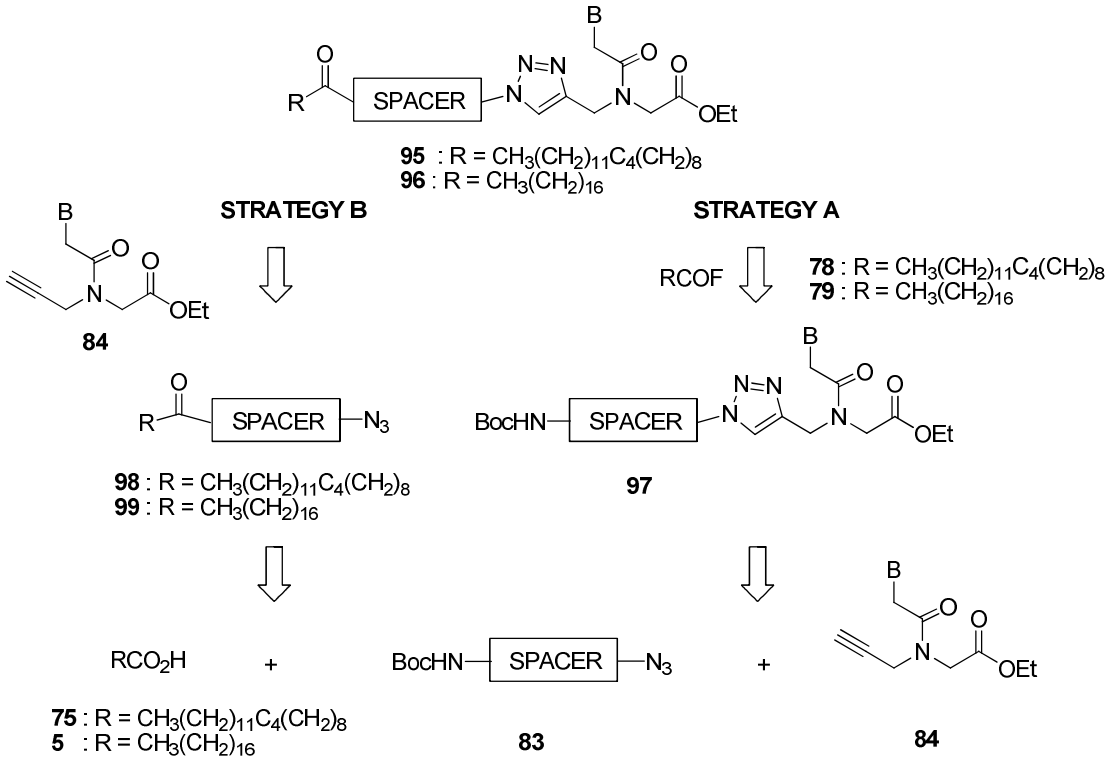
Figure 58 [166]

After the formation of **91**, the next step in the mechanistic pathway involves the distal nitrogen of the azide in **91** attacking the C-2 carbon of the acetylide, to yield the unusual copper(III) metallacycle **92** (step **C**). The calculated energy barrier for this step was found to be 14.9 kcal / mol, which is lower than the barrier reported for the uncatalysed reaction. This discovery can be used to explain the enormous rate acceleration found for the copper(I)-catalyzed process (7 to 8 orders of magnitude) as compared to the uncatalysed thermal cycloaddition.

The final two steps in the reaction mechanism postulated by Sharpless *et al.* involve first ring contraction of **92** to afford the triazolyl-copper derivative **93** (**D** – *Scheme 25*). Subsequently, proteolysis of **93** releases the triazole product **94**, thereby completing the catalytic cycle. Sharpless *et al.* have also shown that this triazolyl-copper intermediate can be captured by electrophiles other than a proton. For example, the incorporation of deuterium at the C-5 of the triazole product has been accomplished in near quantitative yield when the reaction was performed in D_2O.[161]

2.3.4 Proposed synthesis for the new monomer models

From the examples described previously (Section 2.3.2), we reasoned that the copper(I) catalysed "click" reaction could be applied to the preparation of the lipid-spacer-PNA monomer models. It was believed that this method would help to overcome our synthetic difficulties, encountered with the peptide linkage. Therefore, the spacer and the PNA monomer were redesigned in order to incorporate the required terminal alkyne and azido moieties respectively. As mentioned earlier, for synthetic reasons we propose that, the spacer should bear the azido function while the PNA monomer should be functionalised at its *N*-terminus with an alkynyl group. The retrosynthetic analysis of the lipid-spacer-PNA monomer models based on this principle is outlined in *Scheme 26*.



Scheme 26 – B = T or C^{Cbz} or A^{Cbz}

Therefore, we needed to develop suitable synthetic pathways to both the azido spacer **83** and the *N*-alkynyl PNA monomers **84**. As it has been the case for the earlier amide-based models (**71** and **72 –** *Section 2.2.3*), we proposed that the desired lipid-spacer-PNA monomer conjugates **95** and **96** could be assembled *via* two strategies, as outlined in *Scheme 26*. Strategy A involved first the use of the 'Click' reaction to connect azide **83** to alkyne **84** to form the 1,4-disubstituted [1,2,3]-triazole heterocyclic compound **97**. This intermediate would then be functionalised with the chosen lipid tail (i.e. stearic or

diacetylenic) using the acid fluoride procedure reported earlier in *Section 2.2.3*, to yield conjugates **95** and **96**. In the alternative approach of strategy B, the azide spacer **83** would first be functionalised with the desired lipid chain to form compounds **98** and **99**. The free azide of **98** (or **99**) would subsequently be reacted with the *N*-alkynyl PNA monomer **84** using a 'Click' reaction, to afford again the lipid-spacer-PNA monomer models **95** and **96**.

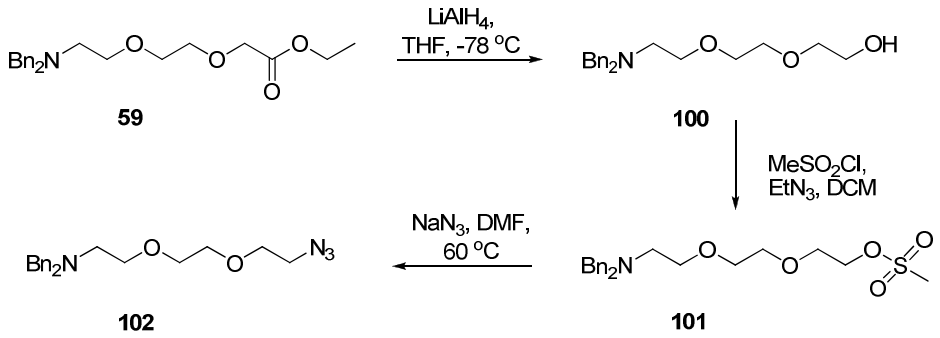
For both strategy A and B, we needed a viable synthetic pathway to AEE azide **83** and the *N*-alkynyl PNA monomer **84**.

2.4 SYNTHESIS OF MODELS USING 1,2,3-TRIAZOLE LINKAGE

2.4.1 Synthesis of the AEEA azide spacer

We decided to start our investigation into the employment of the 'Click' reaction for the preparation of lipid-spacer-PNA monomer conjugates with the synthesis of the Boc-protected AEE azide **83**, since the preparation of the Boc-AEEA analogous **40** (*Section 2.2.1*, page 60) has limited our studies into the amide-linked compounds.

Many reports have been found in the literature for the introduction of azide functions into organic molecules. One of these strategies involves the displacement of a mesylate group with sodium azide. Mesylates are very good leaving groups that can be easily prepared from the corresponding alcohol. In Section 2.2.1 (page 68), we had reported that ethyl ester **59** could be successfully prepared in a reasonable yield in two steps from commercially available 2-(2-aminoethoxy)ethanol **39**. We reasoned that the ester function of **59** could then be reduced to alcohol **100**. This alcohol could subsequently be converted to the mesylate intermediate **101** which could be displaced in the final step to yield azide derivative **102** (*Scheme 27*). Our initial exploration of this synthetic pathway was focussed on model reactions using *N*-protected dibenzyl analogues. However, it was anticipated that this protecting group could be exchanged for the more appropriated Boc analogue once a suitable pathway as been developed, to yield the desired Boc-protected spacer **83**.



Scheme 27

Therefore, the first step involved the reduction of ethyl ester **59** to the corresponding alcohol **100**. Ester reduction is a common reaction in organic synthesis and is usually carried out using hydride reagents such as $LiAlH_4$ or $NaBH_4$. $LiAlH_4$ has been reported to be more powerful than $NaBH_4$ as it can reduce esters under ambient conditions. Therefore, we decided to first explore the use of $LiAlH_4$ in this reaction. Thus, a solution of ester **59** in anhydrous THF at -78 °C was carefully treated with a slurry of $LiAlH_4$ in THF. After aqueous work-up and column chromatography, the alcohol **100** was obtained in 72% yield. Our original method for quenching the reaction, involving addition of a water saturated solution of diethyl ether, limited the scale of the reaction. However, this restriction was eventually overcome by employing another approach for destroying excess $LiAlH_4$. This involved slow addition aqueous 2 M sodium hydroxide to the reaction mixture followed by stirring at room temperature for one hour.[167] The reaction was then filtered through a Celite bed to remove the salts. When adopting this work-up method, the yield of **100** remained unchanged.

Due to this limitation, we decided to investigate the use of $NaBH_4$ in this reduction reaction. Although $NaBH_4$ is not a viable reducing agent under ambient conditions, its reactivity has been found to be greatly improved by the addition of methanol. De Souza *et al.*[168] have reported a simple and efficient route using the $NaBH_4$-MeOH system for the reduction of aliphatic and aromatic methyl and ethyl esters. This approach was therefore attempted for the reduction of ester **59**. A solution of **59** in THF was heated to gentle reflux and then small portions of solid $NaBH_4$ were carefully added. Once the addition was completed, methanol was then very cautiously added. The resulting mixture was kept at reflux until TLC analysis showed that no starting material remained. After work-up and purification by column chromatography, the alcohol **100** was obtained in an 82% yield. In both cases, the signals corresponding to the ester group in **59** had disappeared

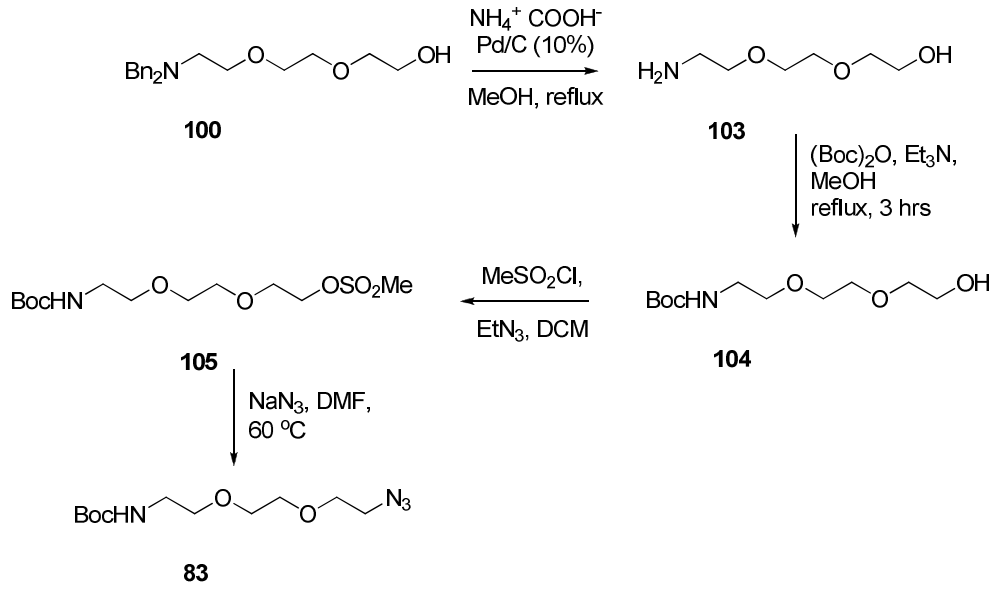
from the ^{1}H and ^{13}C NMR spectra recorded for **100**. With the $NaBH_4$ method, in addition to **100**, another product was isolated in 20% yield and this was identified as being the starting alcohol **42** (*Scheme 12*, page 68). This implied that although ester **59** had been deemed to be pure according to the analytical data collected, a small amount of starting material **42** must have been present.

The next step towards azide **102** involved the preparation of the intermediate mesylate **101** (*Scheme 27*). This was undertaken according to the protocol reported by Smith *et al.*.[169] Thus, triethylamine was added to a solution of alcohol **100** in anhydrous DCM at *ca.* 5 ^{o}C. Subsequently, methanesulfonyl chloride was slowly added and the solution was left to warm up to room temperature. After work-up and purification by column chromatography, mesylate **101** was obtained in an 82% yield. The formation of mesylate **101** was verified by ^{1}H NMR spectroscopy, with the new singlet appearing at 2.92 ppm being assigned to the CH_3 group of the sulfonyl function. The IR spectrum obtained for **98** showed two new strong vibrational bands at 1176 cm^{-1} and 1354 cm^{-1}, corresponding respectively to the symmetric and anti-symmetric S=O stretch of the mesylate group.

The final step in this synthetic pathway involved treatment of a solution of mesylate **101,** in anhydrous DMF, with sodium azide at 60 ^{o}C. The progress of the reaction was followed by TLC analysis until the starting material had been consumed. After work-up and evaporation of the solvent, azide **102** was obtained as a yellow liquid in 87% yield. ^{1}H and ^{13}C NMR spectra recorded for **102** showed the complete disappearance of the signals corresponding to the CH_3 function of the mesylate group. The IR spectrum recorded for **102** showed the disappearance of the vibrational bands for the SO_2 group and the appearance of a new strong band at 2109 cm^{-1} corresponding to the azide group.

Thus, a viable synthetic pathway to dibenzyl-protected azide spacer **102** has been developed starting from the dibenzyl-protected ethyl ester **59**. However, the azide **102** obtained was not suitable for our overall synthetic strategy to lipid-spacer-PNA monomer conjugates. Ideally we required the Boc-protected azide **83**. At this point in this synthetic pathway, we therefore needed to substitute the dibenzyl amino protecting group for Boc. We reasoned that the best place for the removal of the dibenzyl protecting groups was following preparation of the alcohol **100** because, if performed at a later stage, the hydrogenolysis required for their cleavage could affect other functionalities

present in the molecule. Therefore, our adapted strategy for the preparation of the Boc-protected azide **83** is shown in *Scheme 28*.



Scheme 28

As it has been the case for analogue **45** (*Scheme 10*, page 69), we screened a number of different methods that were available for the removal of the dibenzyl protecting groups. The first method involved hydrogenation of **100** in methanol in the presence of Pd/C catalyst at atmospheric pressure. The amino alcohol **103** had been formed in a nearly quantitative yield. This result was surprising as this reaction had proved unsuccessful for the analogous compound **45**. ^{1}H and ^{13}C NMR spectra recorded for **103** confirmed the successful cleavage of the benzyl groups but on the spectra small impurities were present. However, due to the polar character of amino alcohol **103**, it could not be purified by column chromatography and compound **103** was used in the next steps.

We also investigated the catalytic transfer hydrogenation (CTH) approach. Thus, Pd/C catalyst (1 eq. w/w) and ammonium formate (2 eq.) were added to a solution of **100** in methanol. The heterogeneous solution obtained was heated to gentle reflux until TLC analysis showed complete disappearance of starting material **100**. After filtration and evaporation of the solvent, the amino alcohol **103** was obtained as an oily residue in nearly quantitative yield. The CTH method exhibited some advantages compare to the atmospheric pressure hydrogenation strategy mentioned above. CTH was easy to set-up and gave the possibility for scaling up the reaction to *ca.* 2 g of starting material **100**. However, it does suffer one major disadvantage in that it requires large quantity of Pd/C

catalyst (at least one equivalent compare to the dibenzyl substrate). Taddei *et al.*[170] have reported a way for overcoming this limitation; they have carried out the reaction in an unmodified domestic microwave oven (600 W), using only 5% of Pd/C (weight with respect to the substrate), ammonium formate and *iso*propanol. Under these conditions, these researchers observed complete removal of both *N*-Cbz and *N*-benzyl protecting groups from a variety of different substrates within minutes.[170] Unfortunately, we were not able to examine this modification due to time constraints.

A final approach was undertaken for the removal of dibenzyl protecting groups using an H-Cube® instrument that Prof. Adams (at Heriot-Watt University) had on loan for a short period of time. H-Cube® is an analytical instrument developed by ThalesNano Nanotechnology Inc., allowing continuous-flow hydrogenation at high temperature and high pressure. In this instrument, the hydrogen gas is safely generated on demand at high pressure from de-ionized water. The advantage of the technique resides in the fact that hydrogen gas is not stored allowing the use of the instrument on an open bench without any risks. Thus, alcohol **100** was dissolved in absolute ethanol (concentration *ca.* 0.05 M) and passed through the H-Cube® instrument wich was set at 75 ^{o}C with the hydrogen pressure at 70 bar. Again product **103** was obtained in quantitative yield after evaporation of the solvent and it was spectroscopically identical to the material afforded by atmospheric hydrogenation and CTH.

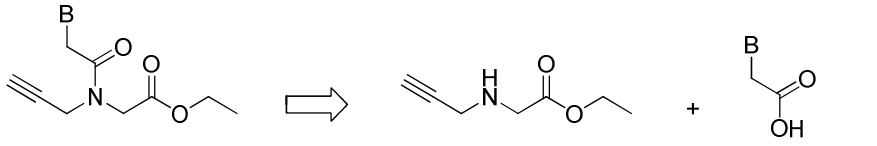
Having successfully removed the dibenzyl protecting group, the free amino group of **103** was re-protected with a Boc group. This was achieved by heating a solution of **103**, di-*tert*-butyl dicarbonate and triethylamine to gentle reflux. After work-up and purification by column chromatography, the *N*-Boc derivative **104** was obtained in an 84% yield.

The last two steps for the preparation of Boc-protected azide **83** involved formation of the mesylate intermediate followed by diplacement of the latter using sodium azide. Both reactions were successfully carried out using the same reagents and conditions as had been used for the synthesis of the analogous dibenzyl azide **102** (*Scheme 27*). The Boc-protected azide **83** was obtained as a yellow liquid in a 71% yield over the two steps.

Thus, we had successfully developed a high yielding synthetic pathway to the desired Boc-protected azide **83** from the dibenzyl-protected ethyl ester **59**. This was ready for use in the 'Click' reaction once the *N*-alkynyl PNA monomers had been prepared.

2.4.2 Preparation of *N*-alkyne PNA monomers

With the synthesis of the Boc-protected azide spacer **83** now complete, we focussed our attention to the synthesis of the novel PNA monomer bearing an alkynyl moiety at its *N*-terminus. It was reasoned that the *N*-alkynyl PNA monomers **84a**, **84b** and **84c** could all be obtained in a similar fashion involving coupling of the appropriate nucleobase acetic acid (**69**, **107**, **108**) to ethyl 2-(prop-2-ynylamino)acetate **106** (*Scheme 29*).



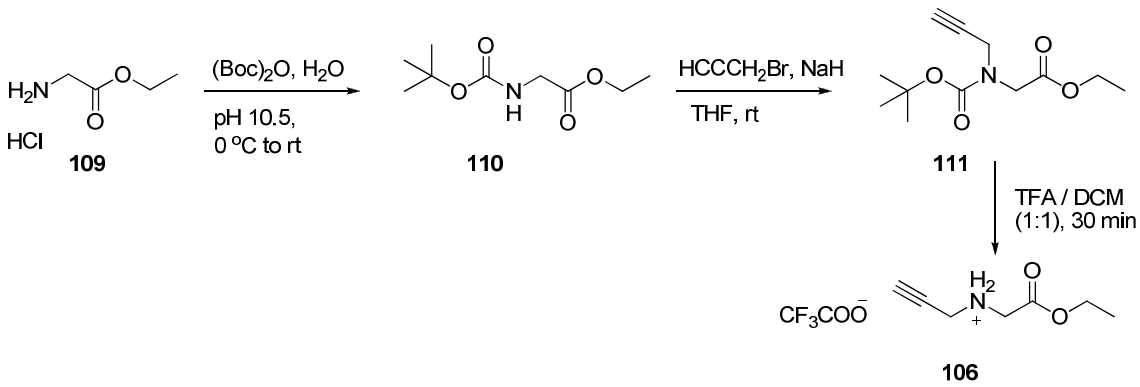
84a : B = T
84b : B = C^{Cbz}
84c : B = A^{Cbz}

106

69 : B = T
107 : B = C^{Cbz}
108 : B = A^{Cbz}

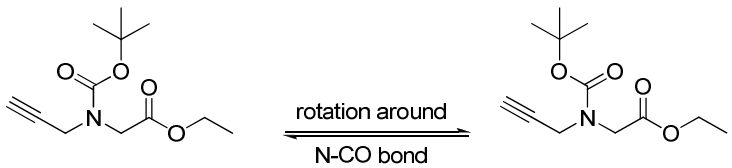
Scheme 29

The novel PNA monomer backbone unit **106** was prepared from commercially available glycine ethyl ester hydrochloride **109** (*Scheme 30*). The amino function of amino acid **109** was firstly Boc-protected by treating a cold aqueous solution of **109** with di-*tert*-butyl dicarbonate whilst keeping the pH of the solution at 10.5. After work-up and column chromatography, the *N*-Boc protected amino acid **110** was obtained in a 63% yield. This reaction could be easily scaled up to *ca.* 3 g of **109** without any effect on the yield being observed.



Scheme 30

In the next step, the nitrogen atom of the carbamate group was alkylated using propargyl bromide and sodium hydride. Hoffmann *et al.*[171] have reported the conditions for a similar reaction and these were employed here. Therefore, a solution of **110** in anhydrous THF was cooled to 0 ^{o}C before being treated with 3-bromopropyne followed by sodium hydride. The reaction mixture obtained was left to stir at room temperature overnight. After careful work-up and purification by column chromatography, compound **111** was obtained in a 50% yield. Along with **111**, a small amount of starting material **110** was also recovered. We investigated whether the yield of **111** could be improved by changing the solvent. Thus, the reaction was repeated in anhydrous acetonitrile and DMF separately. In both cases an increase in the yield of **111** was observed; **111** was produced in 68% yield for the former and 76% for the latter. The ^{1}H NMR spectrum recorded for **111** showed the appearance of a narrow triplet at 2.22 ppm corresponding to the terminal alkynyl proton coupling to the new CH_2 now attached to the carbamate. Interestingly, other signals in the same spectrum appeared doubled; for example, the signal for the NCH_2CO protons was found at 3.89 and 4.09 ppm whereas the signal for the HCCCH_2N protons was found at 4.13 and 4.19 ppm. This splitting was also evident in the ^{13}C spectra and 2D NOE experiments recorded for **111**. From our knowledge of the NMR characteristics of PNA monomers,[139] we hypothesized that this splitting was due to the existence of rotational isomers, arising from the restricted rotation around the N-CO bond as represented *Figure 59*. Such rotamers have been observed earlier in the NMR spectra reported for the thyminyl PNA monomer **70** (Section 2.2.2, page 74).



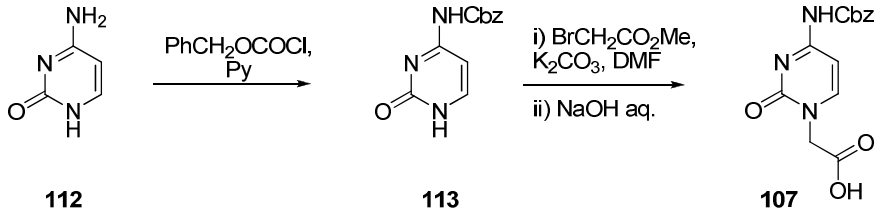
111b - *cis* **111a** - *trans*

Figure 59

The final step in the synthetic pathway to ethyl 2-(prop-2-ynylamino)acetate **106** (*Scheme 30*) involved removal of the Boc group. This was carried out using TFA in dichloromethane to afford the trifluoroacetic acid salt of **106** in quantitative yield. In the ^{1}H NMR spectrum recorded for **111**, the disappearance of the singlets at 1.39 ppm and 1.49 ppm both corresponding to the protons of the Boc group of **111** (both rotamers) confirmed the effective cleavage.

The novel PNA backbone **106** was now ready to be coupled to the appropriate acetic acid derivatives of the nucleobases as previously shown in *Scheme 29*. The preparation of thyminyl acetic acid **69** has already been described earlier in Section 2.2.2. For these studies we also required the preparation of the cytosinyl **107** and adeninyl **108** analogues.

In order to prepare the cytosinyl acetic acid **107**, the exocyclic amino group of cytosine had to be first protected so as to prevent its interference in the later coupling step. A Cbz protecting group was selected as this has been used by Nielsen *et al.*[1] for the preparation of cytosinyl PNA monomer. Therefore, a solution of cytosine **112** in anhydrous pyridine was treated with benzyl chloroformate under the conditions reported by Nielsen *et al.*[1] to yield **113** in a 54% yield (*Scheme 31).* The addition of this lipophilic group was found to improve considerably the solubility of cytosine in organic solvents. Having formed **113**, this compound was alkylated at its N(1) position using methyl bromoacetate in the presence of potassium carbonate. The intermediate methyl ester was not isolated but directly cleaved under basic conditions to afford the desired carboxylic acid **107** in a 30% yield. The yield obtained was consistent with that reported by Nielsen *et al.*.[1]

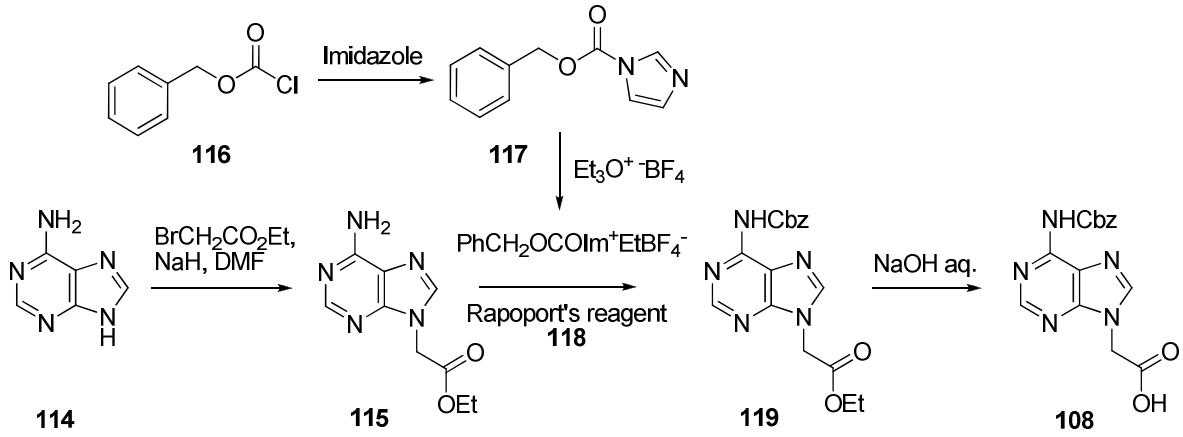


112 **113** **107**

Scheme 31

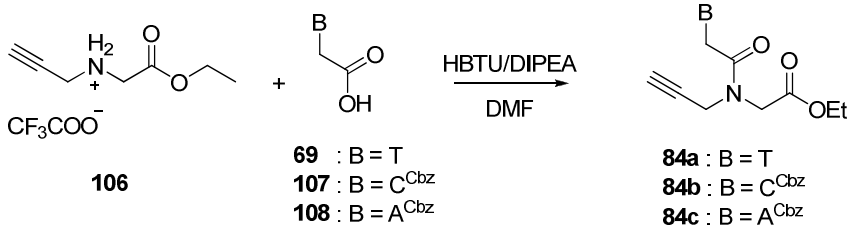
Adenine has an exocyclic amino group that requires protection too. However, in order to ensure that alkylation occurs at the N(9) position the exocyclic amino group of adenine has to be protected after alkylation according to Nielsen *et al.*.[1] Thus, adenine **116** was treated with ethyl bromoacetate in the presence of sodium hydride following the procedure reported by Nielsen *et al.*.[1] This gave **115** in 41% yield after work-up and purification. This procedure yielded only the alkylation at the N(9) position (verified by X-Ray crystallography by Egholm *et al.*[172]). Nielsen *et al.*[1] also reported that **115** gave rise to a complex mixture when reacted with benzyloxycarbonyl chloride (**116**) under a variety of conditions.[1] However, they found that when **115** was reacted with an excess of *N*-Cbz-*N'*-ethylimidazolium tetrafluoroborate **118** (Rapoport's reagent[173, 174]) for 24 hours,

compound **119** could be isolated in good yield. Thus, this procedure was repeated here and ethyl (N^6-Cbz-adenin-9-yl)acetate **119** was obtained in a 90% yield. The final step involved hydrolysis of the ethyl ester of **119** which was successfully achieved by treating a suspension of **119** in methanol with a 2 M aqueous solution of sodium hydroxide, at 0 °C. After work-up and filtration, carboxylic acid **108** was obtained in near quantitative yield (*Scheme 32*).



Scheme 32

With supplies of the thyminyl acetic acid **69**, Cbz-protected cytosinyl acetic acid **107** and Cbz-protected adeninyl acetic acid **108** to hand, we could now examine their coupling to the ethyl 2-(prop-2-ynylamino)acetate **106**. Thus, a solution of **69**, **107** or **108** in DMF was treated with HBTU in the presence of DIPEA. After 15 minutes, this solution was added to a solution of **106** in DMF and DIPEA. After work up and purification by column chromatography, *N*-alkynyl PNA monomers **84a**, **84b** and **84c** were obtained in 60%, 51% and 36% yield, respectively (*Scheme 33*).



Scheme 33

The ^1H and ^{13}C NMR spectra recorded for **84a**, **84b** and **84c** showed 'splitting' of some of the signals as it had been observed for the Boc-protected intermediate **111**. Again this splitting was reasoned to be due to the presence of rotational isomers. By comparing the

integrations of the split peaks in the ^{1}H NMR spectra we deduced that the two rotamers were present in a *ca.* 1:1 ratio. Thus, for these compounds, there did not appear to be any rotameric preference as it had been the case for the Nielsen's PNA monomers. In order to confirm this observation, we decided to carry out NOE experiments on one of these novels PNA monomers. For these studies, we chose to use the cytosinyl PNA monomer **84b**. Based on our knowledge from literature reports we reasoned that the two possible isomers of **84b** are as shown in *Figure 60*; the carbonyl group can either point towards the ester group (**84ba** – *trans* isomer) or towards the propargyl unit (**84bb** – *cis* isomer). The NOE contacts which we were interested in are represented with red arrows.

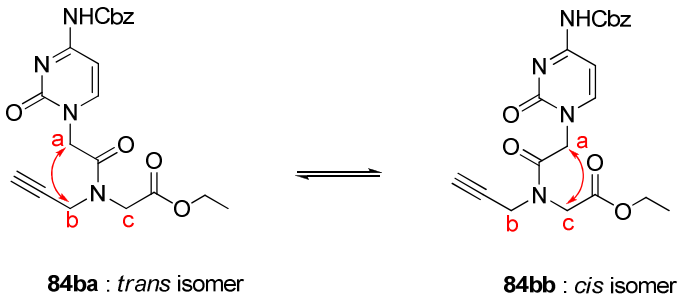


Figure 60 – NOE contacts shown in red

These contacts arose from interactions between the protons of the methylene groups, **b** and **c**, on the backbone unit and the side chain methylene protons next to the nucleobase, **a**. The ROESY* spectrum recorded for **84b** is shown in *Figure 61*. The intensities of the resulting contacts **a—b** (found in the *trans* isomer) and **a—c** (i.e. *cis* isomer) are similar in intensities. These observations confirmed the first deduction obtained from the 1D ^{1}H NMR mentioned earlier, i.e. that there is not rotameric preference in the *N*-alkynyl PNA monomers. This can be understood from the structure of the backbone, as the bulky Boc group of **70** has been replaced with a smaller and linear alkyne group. The rotation around the tertiary amide bond in **84a** is eased due to the absence of the hinderance of the Boc group. However, further complex calculations would have been required to confirm this hypothesis.

* ROESY spectrum obtained by Dr Alan Boyd.

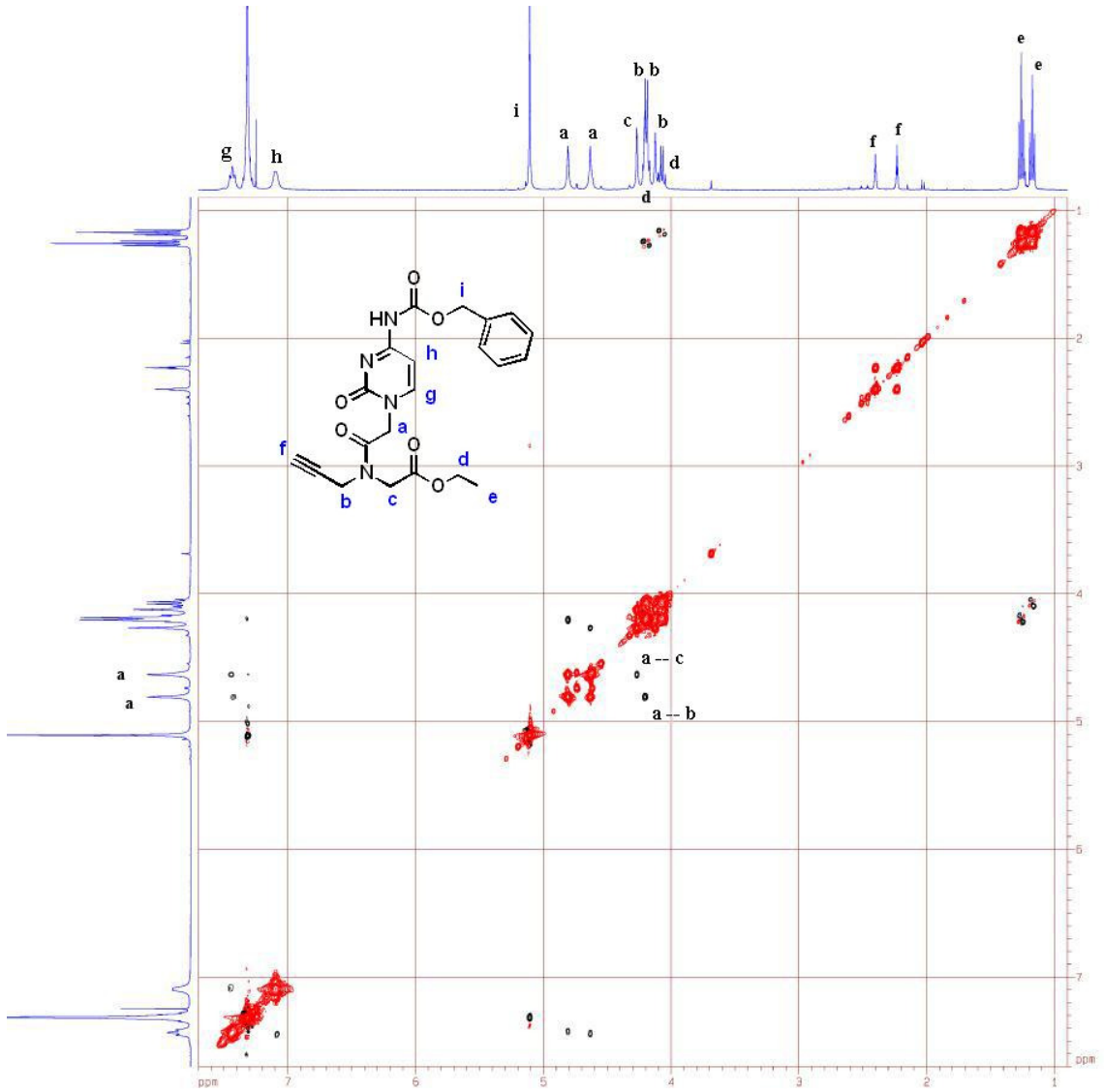


*Figure 61 – ROESY of **84b** ($CDCl_3$, 25 ^oC)*

Subsequently, an X-Ray crystal structure for **84a** was obtained from crystals grown in diethyl ether and methanol using the same slow vapour diffusion method as used for **70**. The ORTEP diagram of **84a** is shown in *Figure 62*. The full crystallographic data of **84a** is available in *Appendix B*. Compound **84a** was found to crystallise in the mono-clinic space group, P2/(1)n. This was the same as it had been observed for thyminyl PNA monomer **70** (Section 2.2.2, page 76).

Although in solution, the tertiary amide bond exists as a rotameric mixture, in crystal structure the amide bond is *trans* (i.e. with the carbonyl pointing toward the glycyl component).

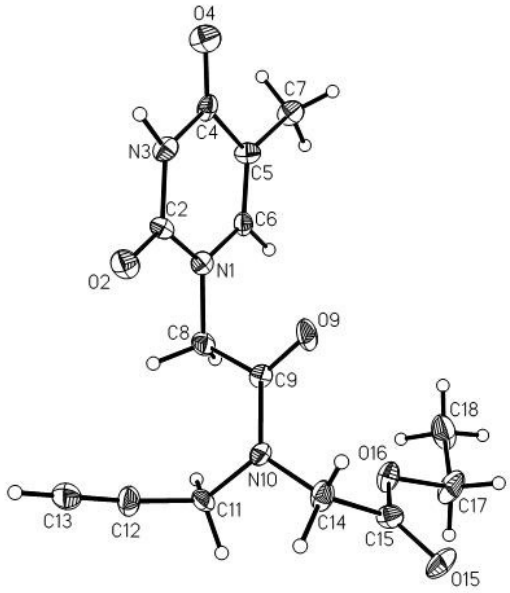


Figure 62

The packing analysis of **84a** (*Figure 63*) showed that it exhibited the same anticipated centrosymmetric hydrogen-bonded dimeric structure with an H---O distance of 1.90 Å as it had been the case for **70**. The hydrogen bond lengths are similar for both cases (c.f. 1.86 Å for **70**).

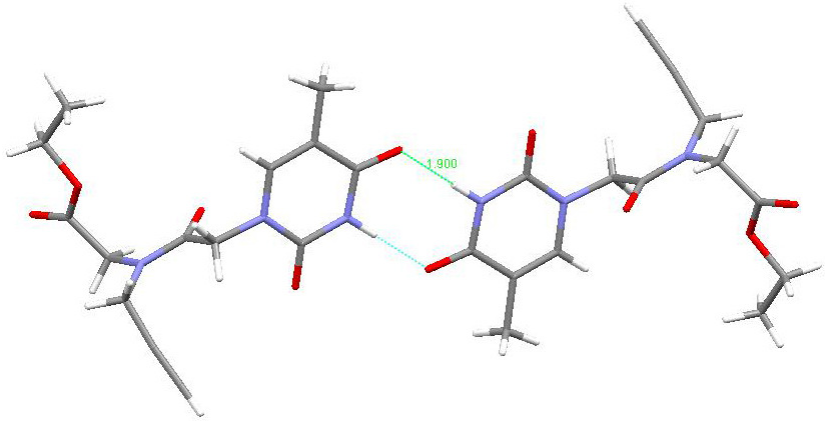


Figure 63

Intermolecular interactions analysis showed that the adjacent molecules of **84a** are held together by a series of donor-H···acceptor type of bonds represented in *Figure 64*. One was between the alkynyl hydrogen H_{13} (*see* ORTEP *Figure 62*) and the oxygen atom of the amide bond (O_{19}) with a H···O distance of 2.30 Å. The oxygen atom of the carbonyl group of the ester was in close contact with the two methylene groups C_{11} and C_{14}. C_{11}-H_{11}···O_9 had a distance of 2.60 Å while for C_{14}-H_{14}···O_9 the H···O distance is 2.46 Å.

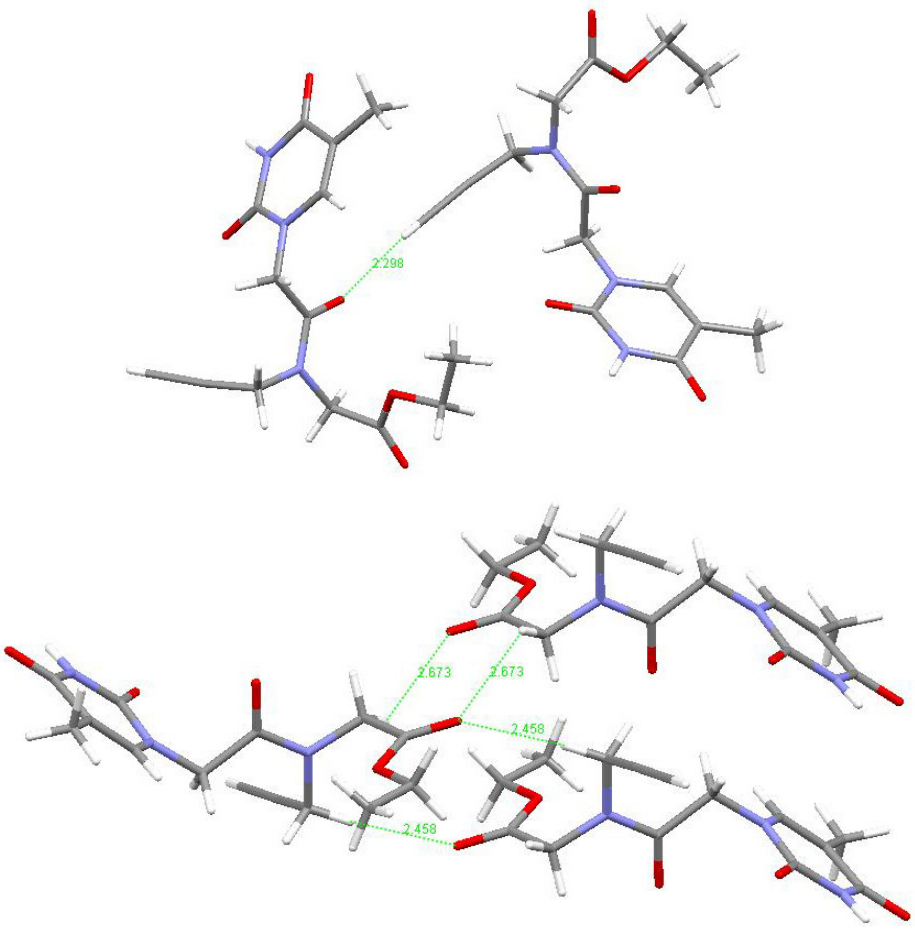


Figure 64

Finally the same kind of π-stacking arrangement of the nucleobases pairs in a stepwise manner is found in **84a** as observed in the case for **70** (*Figure 65*). This time, however, the "steps" were not as long as those found for **70**. This is probably because in **84a** the bulky Boc group has been replaced by the alkynyl group which allows the molecules to lie closer to one another. The vertical distance between two carbonyls of two stacked thymine rings was 4.90 Å.

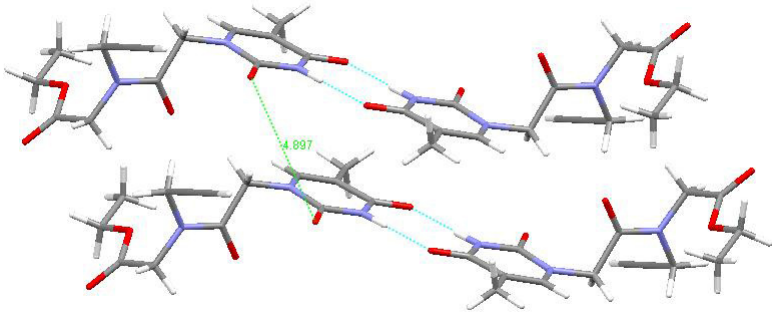


Figure 65

Although the crystal structure of the unmodified PNA monomer **70** has never been reported in the literature, crystal structures of PNA:PNA, PNA:DNA and PNA:RNA duplexes have been described. The X-ray structures of other modified PNA monomers have also been described, e.g. the cyanuryl-PNA monomer[175] and the aminocyclohexyl glycyl PNA monomers[176] (*Figure 66*).

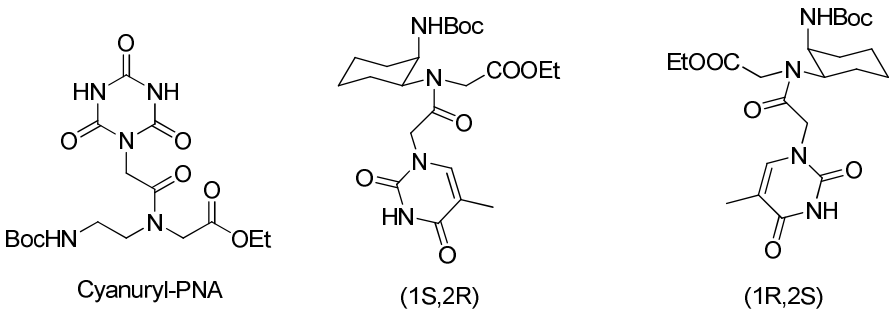


Figure 66

We have calculated the characteristic torsion angles of **70** and **84a** and compared them to those reported in the literature for other PNA structures. These torsion angles are reported in *Table 4* and they are defined as in the report by Nielsen *et al.*[177] as highlighted in *Figure 67*.

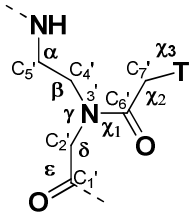


Figure 67

Compound	α	β	γ	δ	χ_1	χ_2
84a	-	43	103	89	-7	173
70	-176	62	68	97	10	178
Cyanuryl-PNA	-77	-60	-86	118	1	142
PNA-PNA[177]						
Strand 1	-112	56	73	114	5	-176
Strand 2	-118	69	65	106	6	-179

Table 4 – Torsion angles (°)[†] in known PNA crystal structures

Torsion angles χ_2 in **84a** and **70** are very close in magnitude to those found in PNA:PNA duplexes. Their values indicate the planarity of the linkage between the nucleobase and the backbone. Ganesh *et al.*[175] proposed that the magnitude of χ_1 can be used to state whether the cyanuryl PNA monomer was locked in the *trans* or in *cis* form. A similar argument can be applied in our case. Although the signs of torsion angle χ_1 are opposite in **70** and **84a**, their relative magnitudes are close and consistent with Ganesh *et al*. This, therefore, confirms that both crystal structures are locked in the *trans* form. In general the values obtained for the torsion angles in **84a** are similar to those found in **70**, apart for torsion angle γ. This departure can be understood by the fact that the bulky Boc substituent in **70** as been replaced by a smaller alkynyl group in **84a**. Therefore, the minimized steric hindrance led to a less restricted torsion angle.

The last investigation carried out on the crystal structures of **70** and **84a** was to check their structural matching and departure.[‡] We decided to set the fitting from the thymine ring to the tertiary amide bond. Unfortunately, the resulting match failed. It was reasoned that this was due to the fact that both compounds have a centrosymmetric space group. This means that the space group has an inversion centre as one of its symmetry elements. Therefore, for every point (x, y, z) in the unit cell, there is an indistinguishable point (-x, -y, -z). Subsequently, we repeated the fitting analysis by inverting one of

[†] Torsion angle definition : $\alpha=C_1'-N-C_5'-C_4'$, $\beta=N-C_5'-C_4'-N_3'$, $\gamma=C_5'-C_4'-N_3'-C_2'$, $\delta=C_4'-N_3'-C_2'-C_1'$, $\chi_1=C_4'-N_3'-C_6'-C_7'$, $\chi_2=N_3'-C_6'-C_7'-N_1$ from *Figure 25*

[‡] Data collected by Dr Georgina Rosair

the molecules to its mirror image. Using this approach we were able to perform structural matching. The result, shown in *Figure 68* and *Figure 69*, was unexpected.

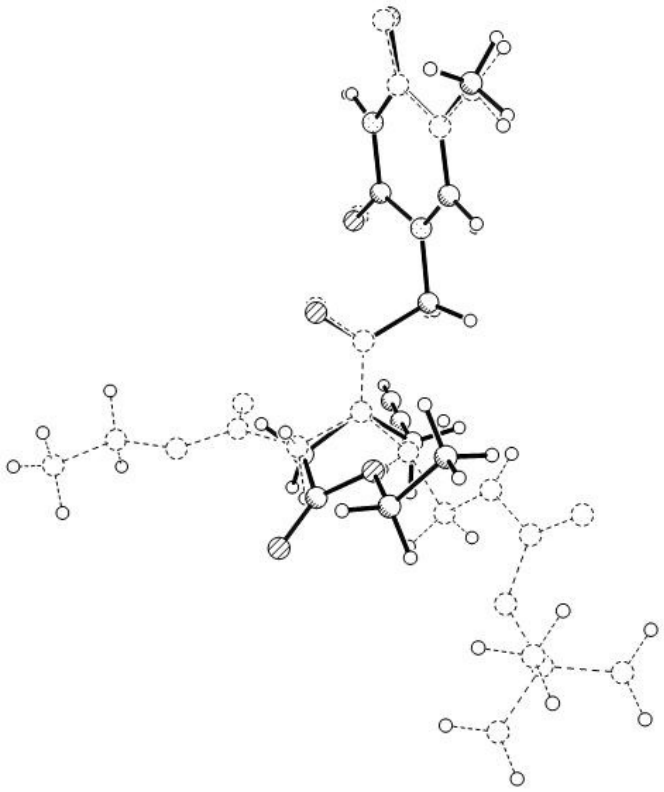


*Figure 68 – **70** in dotted line, **84a** in solid line*

From *Figure 68*, it can clearly been seen that both compounds matched well from the thymine ring down to the tertiary amide bond. However, the glycyl unit of **84a** did not correspond to the one of unmodified PNA monomer **70**. This departure is clearer in *Figure 69* that represents a view in which the thymine rings reside perpendicular to the plan of the page. From this projection it is apparent that there is a difference in the orientation of the CH_2 group joining the ester group to the amide in **70** and in **84a**. This difference in orientation is also found in the methylene group on the other side of the amide. The reasons for these are not fully understood and due to time pressing they were not further investigated.

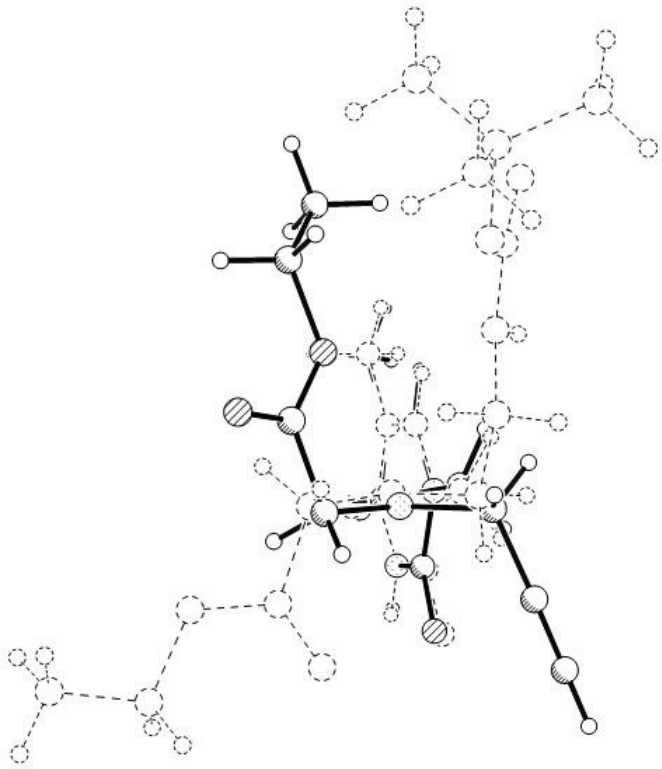
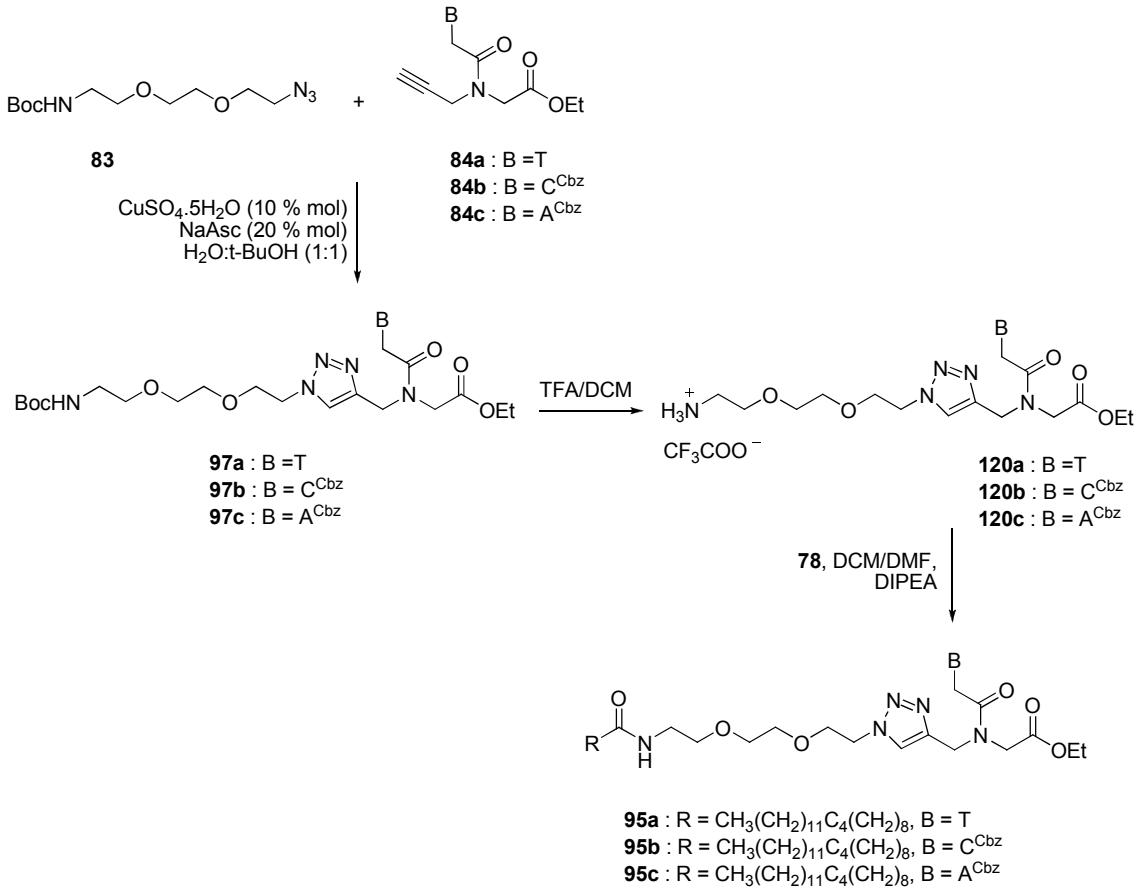


*Figure 69 – **70** in dotted line, **84a** in solid line*

Thus, in conclusion the preparation of the novel *N*-alkynyl PNA monomers bearing thyminyl, cytosyl and adenyl moieties (respectively **84a**, **84b** and **84c**) was successfully accomplished. The focus of this project now moved on the preparation of the lipid-spacer-PNA monomer models. As mentioned in Section 2.3.4, two strategies for their construction were possible and our investigations into each will be discussed in the following two sections (2.4.3 and 2.4.4).

2.4.3 Synthesis of lipid-spacer-PNA monomer models using Strategy A

Our investigation in preparation of the lipid-spacer-PNA monomer models **95** started with the application of the Strategy A, described previously in *Scheme 26* (page 91) and represented below in *Scheme 34*.

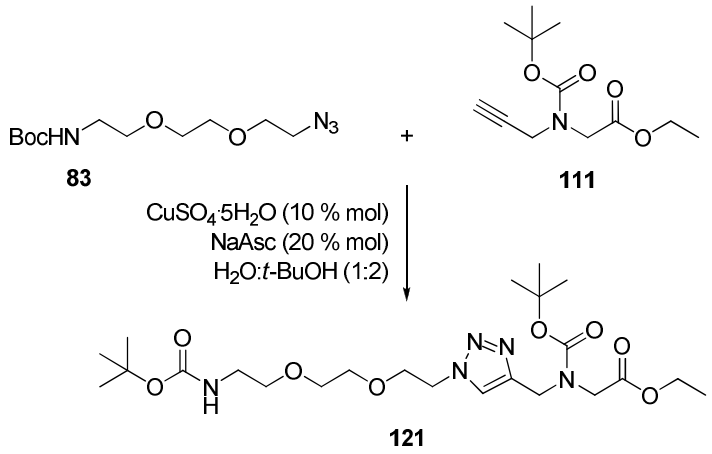


Scheme 34

We were interested in first verifying that the 'Click' reaction could be applied to our own substrates. Before using our PNA monomers (**84a**, **84b** and **84c**), we decided to perform a model reaction using the *N*-alkynyl intermediate **111** (*Scheme 30*, page 97), which contained a Boc group in place of the nucleobase. Once the reaction conditions had been optimised, we proposed to replace intermediate **111** with the *N*-alkynyl PNA monomers (**84a**, **84b** and **84c**).

This model reaction was performed following the procedure reported by Fokin *et al.*.[158] Thus, solid $CuSO_4 \cdot 5\ H_2O$ was added to a solution of azide **83** and alkynyl intermediate

111 in *tert*-BuOH and water (2:1) followed by a freshly prepared aqueous solution of 1 M sodium ascorbate (*Scheme 35*).

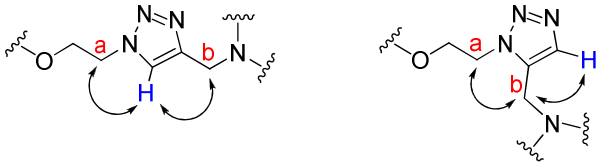


Scheme 35

The colour of the solution obtained quickly changed from colourless to bright yellow as a result of the reduction of the copper metal from Cu(II) to Cu(I). The resulting solution obtained was left to stir at room temperature and the progress of the reaction was followed by TLC analysis. This quickly revealed the formation of a new product along with the gradual disappearance of starting materials. After 24 hours, TLC analysis did not show any further changes. By this time, the solution had become green in colour. This observation implied that Cu(II) had been reformed. Thus the reaction was worked up and the crude mixture obtained was carefully purified by column chromatography. Product **121** was obtained in a 75% yield.

The ^{1}H NMR spectra recorded for **121** showed a combination of all the signals from starting materials **83** and **111**, in the correct integration. The disappearance of the signal for the alkynyl proton of **111** at 2.22 ppm along with the appearance of a new singlet split at 7.56 ppm and 7.75 ppm corresponding to the C(5) triazolyl proton, indicated that [1,2,3] triazole ring had been formed. This was also verified in the ^{13}C NMR spectrum recorded for **121**. As it had been seen before for the PNA monomer, some of the signals from compound **121** were split in both ^{1}H and ^{13}C NMR spectra. Once again this implied two rotamers were present. For example, as just mentioned in the ^{1}H NMR, the signal the C(5) triazolyl proton was found at 7.56 ppm and 7.75 ppm, whereas in the ^{13}C NMR spectra, the signal for the triazolyl C(5) was found at 123.1 ppm and 123.9 ppm.

In order to confirm that the expected 1,4-disubstituted [1,2,3] triazole regioisomer had been formed rather than the 1,5-regioisomer a NOESY ^{1}H NMR of **121** was recorded. It was reasoned that if the 1,4-regioisomer had been formed, the contact between the triazolyl C(5) proton with CH_2 groups on either side (**a** and **b** – *Figure 70*) would be observed. However, if the 1,5-regioisomer was present, the triazolyl C(5) proton should only show contact with the CH_2 **b**. CH_2 **b** would also show a contact with CH_2 **a** in the latter regioisomer.



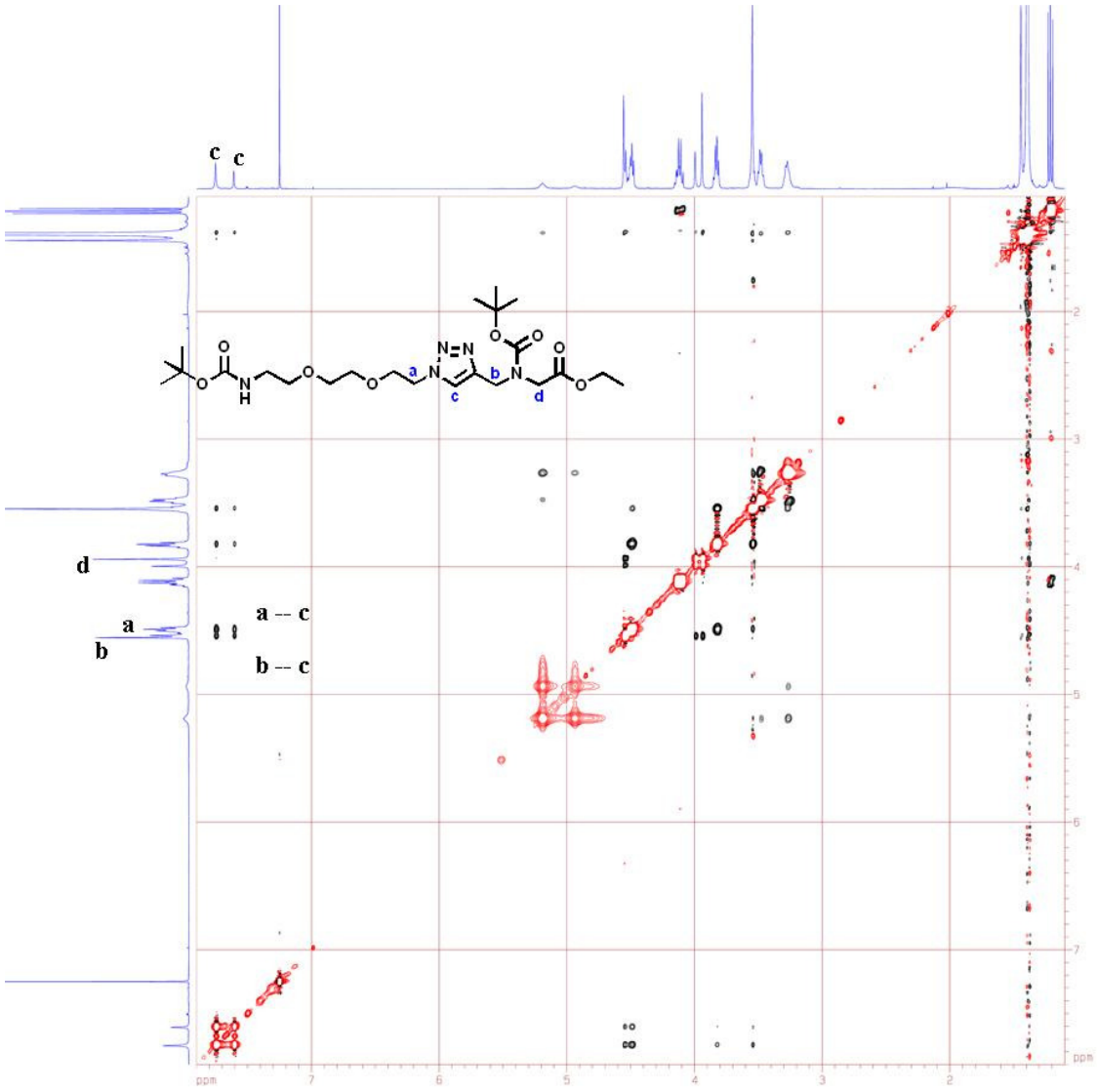


*Figure 70 –NOESY of **121** (400 MHz, $CDCl_3$, 25 ^{o}C)*

The NOESY ^{1}H NMR spectrum recorded for **121** is shown in *Figure 70*. This clearly shows the predicted characteristic NOE contacts for the 1,4-regioisomer. No contacts corresponding to the 1,5-regioisomer were evident. This experiment confirmed that this 'Click' reaction had given rise to the 1,4-disubstituted [1,2,3]-triazole only as expected.

Following the successful coupling of azide **83** to intermediate **111** using the 'Click' reaction (*Scheme 35*), we decided to repeat this reaction exchanging the model alkynyl **111** for *N*-alkynyl PNA monomers (**84a**, **84b** and **84c –** *Scheme 34*). Initially, we worked with the thyminyl PNA monomer **84a**. When we performed the 'Click' reaction with **84a**, we modified the procedure slightly from the conditions used for the model alkyne **111**. This was because we noticed that the PNA monomer **84a** did not readily dissolve in the original solvent mixture used (i.e. *tert*-butanol and water). This limitation was overcome by first dissolving the *N*-alkynyl PNA monomer in dichloromethane. This solution was then added to the aqueous solution containing the azide **83**, closely followed by the copper catalyst and the solution of sodium ascorbate. This modification also simplified work-up of the reaction, as a simple separation of the aqueous phase from the organic phase enabled the organic reaction products to be easily separated from the copper salts. The crude product afforded was subsequently purified by column chromatography to afford **97a** in 72% yield. The spectral analysis of ^{1}H and ^{13}C NMR showed all the signals expected for **97a**. Again some signals were split due to the presence of rotamers. This reaction was repeated using N^{4}-Cbz cytosinyl PNA monomer **84b** and N^{9}-Cbz adeninyl PNA monomer **84c** in the place of **84a** to yield the corresponding intermediates **97b** (76%) and **97c** (64%) respectively (*Scheme 34*).

The next step in Strategy A, for the preparation of 'lipid-spacer-PNA monomer' model compounds, involved connection of the lipid chain to the *N*-terminal of the intermediates **97a**, **97b** and **97c**. This connection involved the formation of a peptide bond, a similar methodology to the one described for the preparation of lipid-spacer intermediates **73** and **74** (Section 2.2.3, *Scheme 17*, page 79) was applied here. In order to carry out this connection, the Boc protecting group of compounds **97a**, **97b** and **97c** has to be first removed. Therefore, starting with the thyminyl intermediate **97a**, its *N*-Boc protecting group was cleaved by stirring a solution of **97a** in dichloromethane in the presence of TFA at room temperature (*Scheme 34*). After one hour, TLC analysis showed that the starting material had been consumed. After work up, analogue **120a** was obtained as its trifluoroacetic acid salt in quantitative yield. The disappearance of the signal at

1.42 ppm corresponding to the Boc group in the ^{1}H NMR spectrum recorded for **120a** indicated that successful *N*-deprotection had been accomplished. Intermediates **97b** and **97c** were similarly deprotected and comparable yields had been accomplished. The ^{1}H NMR spectral analysis showed similar features as for **120a**.

The trifluoroacetic acid salts obtained were used without further purification in the coupling step, which involved reaction with the appropriate lipid fluorides. The preparation of diacetylenic and stearoyl acid fluorides **78** and **79** has been reported in Section 2.2.3 (page 79). The reaction between the deprotected *N*-terminus of **97a**, **97b** and **97c** and the acid fluoride should afford the 'lipid-spacer-PNA monomer' models. Thus, the thyminyl derivative **120a** was treated with diacetylene acid fluoride **78** under the same conditions as described earlier in Section 2.2.3 (page 80). After work-up and purification by column chromatography, conjugate **95a** was obtained in a 71% yield. The ^{1}H NMR spectrum showed all the signals corresponding to the three building blocks in their ratios. For example, the CH_3 group of the diacetylene chain and the thyminyl C(5)CH_3 were found at 0.88 ppm and 1.90 ppm, respectively. Again, some signals were split, indicating the presence of two rotational isomers. These were estimated to be in a 1:1 ratio. HRMS supported the preparation of **95a** where the $[M+H]^+$ molecular ion was found at 838.5443. Theoretically, $C_{45}H_{72}N_7O_8$ requires a m/z of 838.5437. This reaction was then repeated using the cytosinyl and adeninyl analogues, **120b** and **120c** respectively (*Scheme 34*) to give the corresponding models **95b** and **95c**. These too were obtained in good yields (70% and 74% respectively). The ^{1}H and ^{13}C NMR spectra recorded were consistent with the preparation of **95b** and **95c**.

A characteristic feature of diacetylenic models **95a-c** is their gradual change on exposure to light from colourless to deep blue in the solid state as a result of a topochemical polymerisation. The same compounds, however, did not seem to polymerise in solution (e.g. they remained stable in the NMR sample in $CDCl_3$).

Even though strategy A worked well for the construction of the 'lipid-spacer-PNA monomer' models we devised to explore strategy B to see if any improvement in yields could be achieved.

2.4.4 Preparation of monomer models based on Strategy B

Strategy A allowed the preparation of the second generation of diacetylenic models bearing thyminyl, cytosinyl or adeninyl functions in good yields (*ca.* 70%). However, the retrosynthetic analysis of preparation of these models showed that another strategy (i.e. strategy B, *Scheme 26*, page 91) could be also used. As mentioned above we were interested in verifying the applicability of strategy B to prepare models **95a-c**.

Strategy B first involved the coupling of the lipid chain (i.e. diacetylenic or saturated) to the *N*-terminal of the azide spacer **83**. Subsequently, the free azide of this intermediate was reacted with the *N*-alkynyl function of the PNA monomers **84a-c**, to afford the 'lipid-spacer-PNA monomer' models desired. A variety of model compounds could be easily prepared by applying strategy B to parallel synthesis. In fact, it was reasoned that the lipid-spacer intermediates could be reacted with different *N*-alkynyl analogues e.g. PNA monomer, dimers and even oligomers (*Figure 71*). Therefore, specificity is introduced to the model compounds in the last step of the synthetic pathway rather than in the first step as it was the case using strategy A.

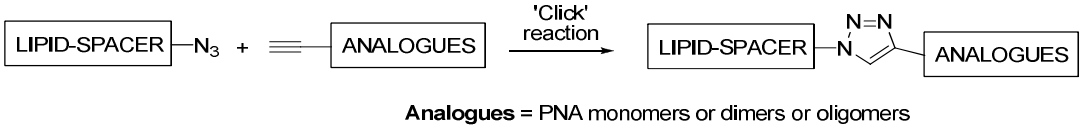
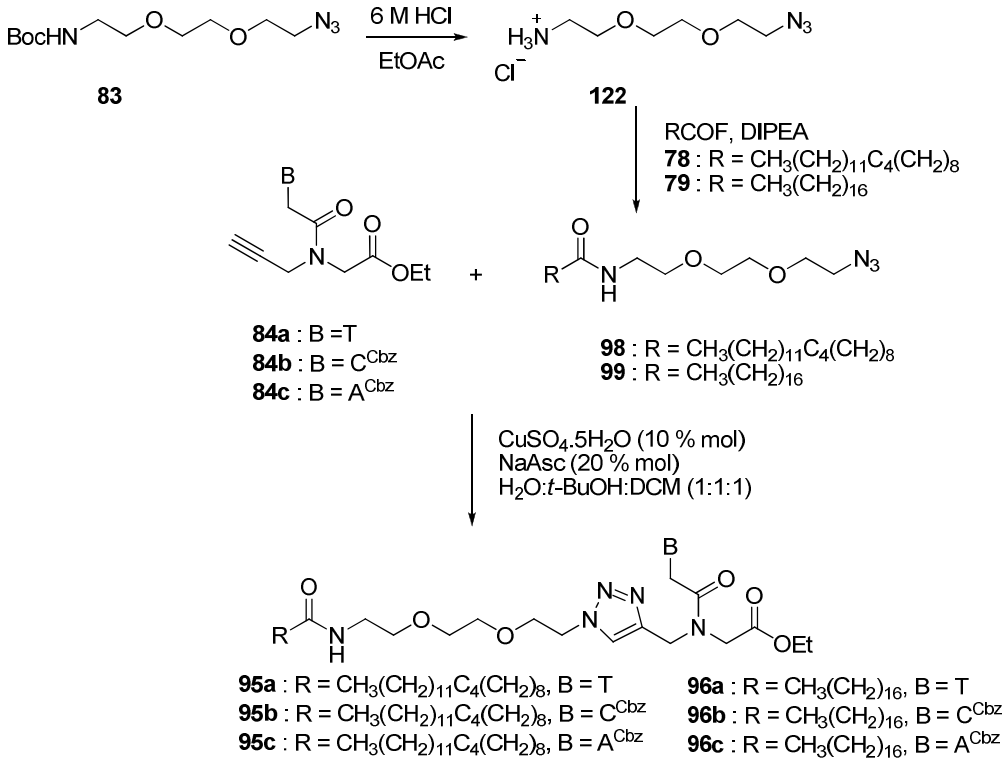


Figure 71

Therefore, the first step in this new synthetic pathway leading to the coupling between the chosen lipid and the *N*-terminus of azide spacer **83** required first the cleavage of the Boc protecting group from **83** (*Scheme 36*). This was achieved by treating a solution of **83** in ethyl acetate with concentrated hydrochloric acid at room temperature. After one hour, TLC analysis showed that the starting material had been consumed. After work up, analogue **122** was obtained as its hydrochloride salt, in a quantitative yield. The disappearance of the signal at 1.42 ppm corresponding to the *tert*-butyl protons in the ^{1}H NMR spectrum recorded for **122** indicated that successful *N*-deprotection had been accomplished.

Subsequently, the *N*-terminus of **122** was coupled to the acid fluoride derivative of 10,12-PCDA **78** under the same conditions as described previously in Section 2.2.3

(page 80). After work-up and purification by column chromatography, the corresponding lipid-spacer intermediate **98** was obtained in 80% yield (*Scheme 36*). The ^{1}H NMR spectra recorded for **98** showed all the signals corresponding to both starting materials in their correct ratio. The appearance of a new broad triplet at 5.55 ppm was assigned to NH proton coupling to the neighbouring CH_2, indicating the formation of the peptide bond. This reaction was repeated using the acid fluoride derivative of stearic acid **79**. After work-up and purification by column chromatography, intermediate **99** was obtained in 80% yield. Again, the ^{1}H NMR spectra recorded for **99** showed all the expected signals in their correct ratio. The formation of the peptide bond in **99** was evident from the appearance of a similar broad triplet at 5.55 ppm (i.e. assigned to NH proton coupling to the neighbouring CH_2) as it had been the case in **98**.



Scheme 36

The saturated compound **99** was found to be stable at ambient conditions and could be stored at -20 ^{o}C for long periods of time without degradation occurring. On the other hand, the unsaturated compound **98** was found to be unstable in the solid state. **98** was obtained as a white solid that very quickly turned blue on exposure to light. This characteristic had already been observed in our earlier work with the diacetylenic models **95a**, **95b** and **95c** (page 113). The change of colour is an indication of the reaction of polymerisation taking place. This change of colour was also obtained when **98** was stored in

the dark at -20 ^{o}C but it did not seem to polymerise in solution (i.e. in the NMR sample in $CDCl_3$). Due to this problem, we reasoned that strategy B was not suitable for the preparation of 'lipid-spacer-PNA monomer' models bearing the diacetylenic lipid tail.

Despite the problems reported above, having prepared **98**, its coupling with the *N*-alkynyl PNA monomers **95a-c** was still investigated (*Scheme 32*). We first selected for this purpose the thyminyl PNA monomer **84a** which was reacted with **98** under the same conditions of 'Click' reaction as described in strategy A (page 112). The *N*-alkynyl PNA monomer **84a** and the lipid intermediate **98** were first dissolved in dichloromethane, followed by the aqueous mixture of water and *tert*butanol (1:1) containing the copper catalyst and the solution of sodium ascorbate. The progress of the reaction was carefully followed by TLC and after 30 hours at room temperature, it showed that the two starting materials **84a** and **98** were still present and that along with them a novel compound had formed. The reaction mixture was further stirred (total time 48 hours) but unfortunately TLC analysis kept revealing the presence of both starting materials. Nevertheless, after this time, the reaction was worked up and, after careful purification by column chromatography, the model **95a** was obtained in 59% yield. This reaction product was spectroscopically identical to the one obtained with strategy A. The synthetic pathway was then repeated for the cytosinyl PNA monomer **84b** and the analogous model **95b** was obtained in a 30% yield. The reaction product was spectroscopically identical to the **95b** obtained in strategy A.

Both yields of **95a** and **95b** obtained using strategy B were lower than when following strategy A. One reason for this is possibly because of the polymerisation problems of starting material **98**. Due to these low yields, we decided not to prepare the analogous adeninyl model **95c** according to this strategy.

With the preparation of the diacetylenic models now completed we decided to move to the synthesis of the models functionalised with a saturated lipid chain. As already discussed in Section 1.5.2 (page 41), Charych *et al.*[98, 103] have demonstrated that when constructing liposomes for use as biosensors it is not always necessary for the bioactive head-group to be covalently linked to the polydiacetylene matrix. Given the stability of intermediate **99** we decided to explore the 'Click' reaction using this starting material. Therefore, 'lipid-spacer' intermediate **99** was reacted with *N*-alkynyl PNA monomers **84a**, **84b** and **84c** separately, under the same reaction conditions described for the con-

nection of the diacetylenic lipid **98**. Once again, after 48 hours of stirring at room temperature, TLC analysis indicated that the reaction did not reach completion. However, after this time, we decided to work up the reaction. After purification by column chromatography, the corresponding models **96a**, **96b** and **96c** were obtained, in 74%, 62% and 55% yields, respectively. The ^{1}H NMR spectra recorded for these models showed the expected combination of all the signals from the starting material **99** and the *N*-alkynyl PNA monomer (**84a-c**), in the correct ratio. For example, in the spectrum of model **96a**, the disappearance of the signal at 2.22 ppm corresponding to the alkynyl proton of **84a** along with the appearance of the new singlet split at 7.71 ppm and 7.94 ppm corresponding to the C(5) triazolyl proton, indicated that [1,2,3] triazole ring had formed. ^{13}C NMR spectra also verified the formation of this ring with the appearance of two new signals, the triazolyl quaternary carbon C(4) was found at 142.4 ppm and 142.8 ppm, and the triazolyl CH group found at 123.8 ppm and 124.2 ppm. As it had been the case for diacetylenic models **95a-c**, some of the signals of the stearoyl models **96a-c** were split in both ^{1}H and ^{13}C NMR spectra, implying the presence of rotameric isomers.

At this stage of the research project we were in the possession of a library of six models made from two different types of lipid chains and three different nucleobases (*Scheme 36*). Two strategies had been carried out to produce these conjugates and from our studies it appeared that in the case of the diacetylenic models, the lipid chain should be connected in the last step of the synthetic pathway. Although we did not attempt strategy A for the preparation of saturated models, the yields obtained with strategy B were really good (above 50%), indicating that this strategy is suitable. We believe that similar yields would have been obtained with strategy A but due to time constraints we did not investigate further.

It is worth noting that the conditions used for the 'Click' reaction were not optimised and that further investigations into their improvement would be necessary in the future. Improvements in yields of 'Click' reaction are regularly being published in the literature. For example, it has been reported by Fokin *et al.*[178] that the use of ligands such as TBTA (*Figure 72*) can improve both the purity and yields of the final product by stabilizing the copper(I) and protecting it from oxidation and disproportionation.

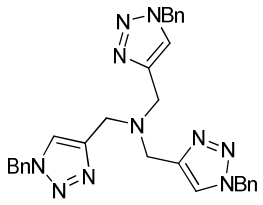
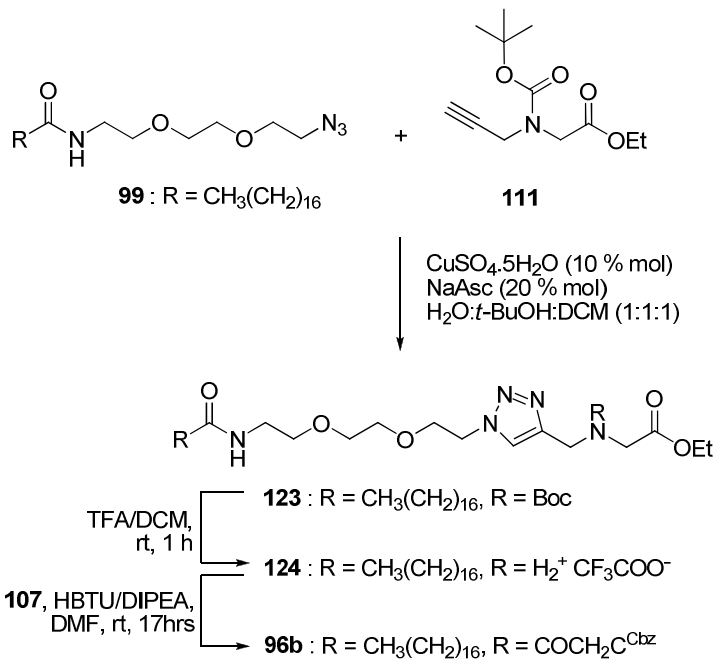


Figure 72 – Structure of TBTA

2.4.5 Preparation of cytosinyl model using strategy C

Finally, during the course of our investigations into strategy B a third strategy was designed. Strategy C was based on the same principle as strategy B but here, the *N*-alkynyl PNA monomer was replaced with the Boc intermediate **111**. This strategy involved the preparation of the lipid-spacer-PNA monomer backbone model **120** that would then be functionalised with appropriate nucleobase acetic acid derivative. Thus, on the same principle as strategy B had been designed (*Figure 71*, page 114), we reasoned that specificity into the models would be inserted in a later step, after the preparation of the backbone of the models.

Strategy C was investigated for the preparation of model **96b** and started with the 'Click' reaction between lipid-spacer intermediate **99** and *N*-alkynyl PNA intermediate **111** (*Scheme 37*).



Scheme 37

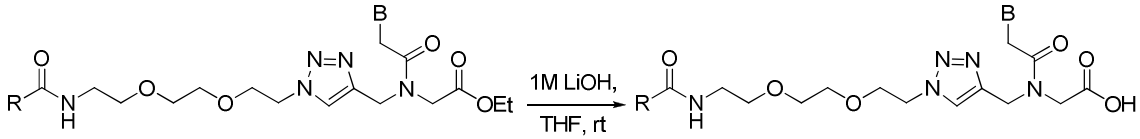
The same conditions described in Section 2.4.4 (page 113) were used here. After work-up and purification by column chromatography, the lipid-spacer-Boc monomer **123** was obtained in an 89% yield. The ^{1}H NMR spectrum recorded for **123** showed all signals expected such as the split singlet at 7.61 ppm and 7.78 ppm corresponding to the triazolyl C(5)H newly formed. ^{13}C NMR analysis also confirmed the formation of the [1,2,3]-triazole ring. Subsequently, the Boc protecting group of **123** was cleaved to liberate the secondary amino group under acidic conditions. The disappearance of the split signal at 1.35 ppm and 1.40 ppm corresponding to the *tert*-butyl protons in the ^{1}H NMR spectrum, confirmed the successful deprotection. The trifluoroacetic acid salt **123** was then directly reacted with cytosinyl acetic acid **107** using HBTU as the activating agent in the presence on DIPEA. After work-up and purification by column chromatography, the 'lipid-spacer-PNA monomer' **96b** was obtained in a 53% yield. The product was spectroscopically identical to the one obtained from strategy B. Although the yield obtained with this strategy is slightly lower than with strategy B (c.f. 62%, page 116), the overall synthetic pathway was simplified since it did not require the preparation of the *N*-alkynyl PNA monomer **84b**. Overall, fewer purification steps by column chromatography were required. However, due to time constraint only the cytosinyl model **91b** has been prepared using this pathway but future investigations involving the other nucleobases acetic acid derivatives should be carried out.

We mentioned earlier that strategy B was not suitable for the preparation of diacetylenic models **95a-c** due to the instability of intermediate **98**. Therefore, it would be interesting to verify the stability of the diacetylenic analogues of **123** and its applicability in this strategy.

2.4.6 Preparation of the corresponding acid models

In the preliminary studies carried out within our research group, Murray[2] discovered that for successful incorporation of the *N*-lipid functionalised PNA models into PDA-liposomes a *C*-terminus carboxylic acid was required. This was evidences with the incorporation of 100% of their 'lipid-PNA monomer' models into PDA-liposomes; the *C*-terminus ester model precipitated during the step of polymerisation leading to pale blue solutions, while the carboxylic acid terminated model yielded dark purple solution after photopolymerisation.

With these results in mind, we decided to also prepare the carboxylic acid analogues of the 'lipid-spacer-PNA monomer' models obtained in sections 2.4.3 and 2.4.4 (**95a-c** and **96a-c**) as shown in *Scheme 38*.



95a : R = $CH_3(CH_2)_{11}C_4(CH_2)_8$, B = T
95b : R = $CH_3(CH_2)_{11}C_4(CH_2)_8$, B = C^{Cbz}
95c : R = $CH_3(CH_2)_{11}C_4(CH_2)_8$, B = A^{Cbz}

96a : R = $CH_3(CH_2)_{16}$, B = T
96b : R = $CH_3(CH_2)_{16}$, B = CCbz
96c : R = $CH_3(CH_2)_{16}$, B = ACbz

125a : R = $CH_3(CH_2)_{11}C_4(CH_2)_8$, B = T
125b : R = $CH_3(CH_2)_{11}C_4(CH_2)_8$, B = C^{Cbz}
125c : R = $CH_3(CH_2)_{11}C_4(CH_2)_8$, B = A^{Cbz}

126a : R = $CH_3(CH_2)_{16}$, B = T
126b : R = $CH_3(CH_2)_{16}$, B = CCbz
126c : R = $CH_3(CH_2)_{16}$, B = ACbz

Scheme 38

Therefore, our investigations into the hydrolysis of 'lipid-spacer-PNA monomer' models started with the thyminyl models **95a** and **96a**. A solution of **95a** was dissolved in THF and treated with 1 M aqueous solution of lithium hydroxide (*Scheme 38*). After work-up and purification by column chromatography, the corresponding acid **125a** was obtained in an 82% yield. ^{1}H NMR spectrum recorded for **125a** evidenced the cleavage of the ester with the disappearance of the signal at 4.21 ppm corresponding to the CH_2 group of the ethyl function of **95a**. In the ^{1}H NMR of **95a**, the signal corresponding to the CH_3 group was hidden under the peak corresponding to the methylene of the lipid chain. ^{13}C spectrum also confirmed the hydrolysis of the ester and the formation of **125a**. Finally, a HRMS further confirmed the formation of **125** with the molecular ion giving rise to a peak at m/z of 810.5113; a compound with molecular formula $C_{43}H_{68}N_7O_8$ is theoretically expected to produce an *m/z* of 810.5124. Similar results were obtained when the saturated conjugate **96a** was treated under the same conditions, to yield **126a** in an identical yield.

After these successful conversions, the cytosinyl model compounds **95b** and **96b** were reacted using the same conditions. Unfortunately, despite numerous attempts to prepare and isolate the corresponding acids (**125b** and **126b**) from the complicated crude reaction mixtures, **125b** and **126b** were not obtained. The reasons for this are still not understood as the reactions conditions employed here were the same as those used earlier by

Murray.[2] Due to the lack of time and material, further investigations were not attempted and our intention to prepare **125b** and **126b** was abandoned.

Lastly, we focused on the hydrolysis of adeninyl models **95c** and **96c**. This time, the conditions were changed and we decided to use a diluted aqueous solution of sodium hydroxide (0.77 M) as reported by Howarth *et al.*[179] for similar purpose. After work up and purification by column chromatography, the corresponding acids **125c** and **126c** were obtained in 92% and 53% yields, respectively. Again the disappearance of the signals for the ethyl group in both ^{1}H and ^{13}C NMR spectra confirmed the cleavage of the ester.

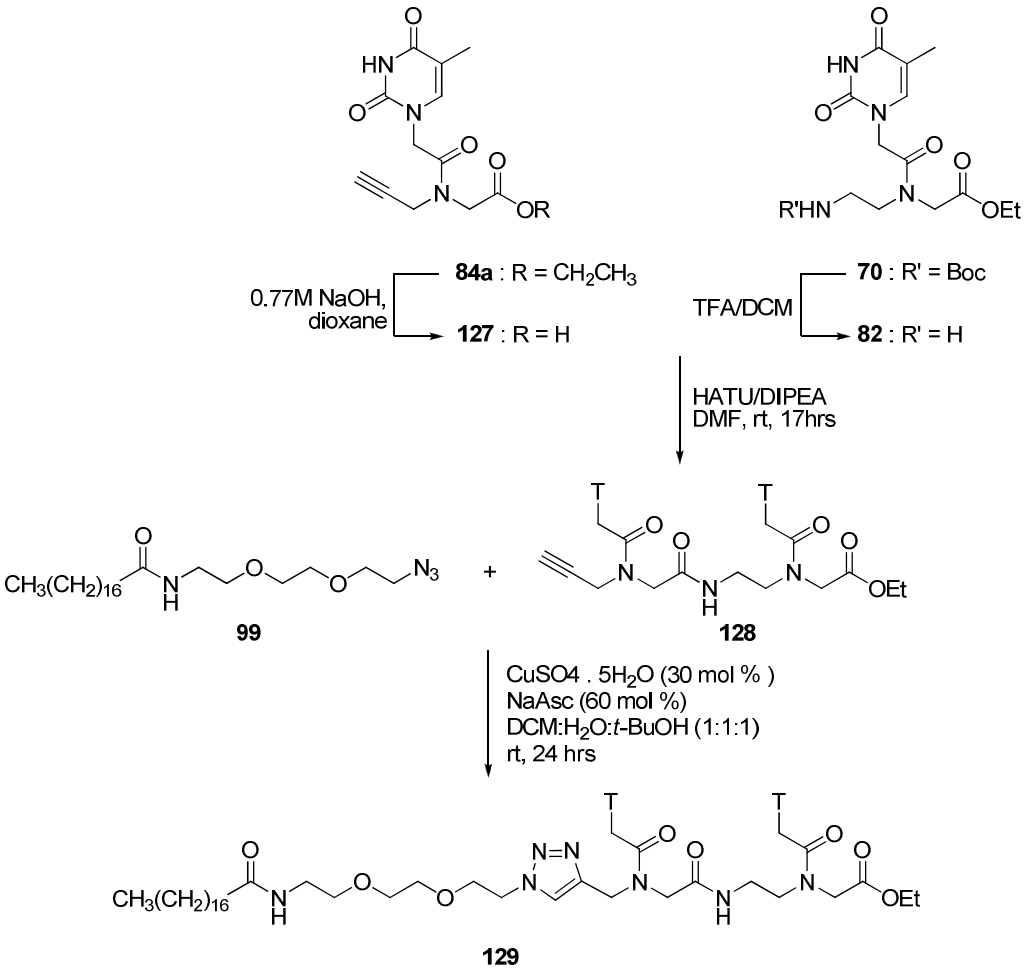
In summary, we now have successfully prepared ten 'lipid-spacer-PNA monomer' models out of the twelve expected. These were prepared from a combination of three different PNA monomers and two types of lipid chain. Six of them had an ester function at their *C*-terminus while the thyminyl and adeninyl have also been prepared in their *N*-terminus carboxylic acid derivatives. We, therefore, decided to move onto the investigation of their abilities to incorporate into PDA-liposomes. This area of the project will be the focus of Chapter 3.

2.4.7 Synthesis of dimer models

It was evidenced that monomer models could be successfully synthesised and the ultimate aim of the synthetic part of this project concerned the preparation of compounds bearing a PNA oligomer as the headgroup. As mentioned earlier (Chapter 1, Section 1.6.2, page 55), Murray[2] reported the preparation of homothymine (T_{10}) oligomers using the Merrifield solid phase approach and following the procedure reported by Christensen *et al.*.[120] Their functionalisation with both diacetylenic and stearoyl lipids was also described by Murray. In our case, the preparation of similar compounds containing the spacer was necessary to fully assess the benefits of this spacer. However, the time spent on model reactions prevented us to perform solid phase synthesis, even though the protocol had been developed in our laboratories. It was still reasoned that we should try to produce models bearing a larger PNA headgroup than a simple monomer. This would give us some insight into the applicability of our synthetic strategies to prepare conjugates bearing a PNA oligomer as headgroup. Therefore, we decided to prepare a thyminyl PNA dimer functionalised with an alkynyl moiety at its *N*-terminus using so-

lution phase synthesis. Then, according to the principle developed in Strategy B (Section 2.4.4, page 114) this PNA dimer would be functionalised with the stearoyl lipid chain. In fact, the polymerisation of lipid-spacer intermediate **98** (page 115) restrained us from using it. We also reasoned that for preliminary explorations we should start with the simplest lipid chain, i.e. stearic acid derivative. Once the synthetic pathway to the thyminyl PNA dimer was optimised, it would be easy to explore its functionalisation with either lipid chains.

Therefore, according to Strategy B, the synthesis of 'lipid-spacer-PNA thyminyl dimer' model **128** involved performing the 'Click' reaction between the azide moiety of the lipid-spacer intermediate **99** and the *N*-alkynyl function of the dimer **128**. We proposed that dimer **128** could be prepared by conducting a peptide coupling between the *N*-alkynyl thyminyl PNA monomer **84a** and the thyminyl PNA monomer **70** (*Scheme 39*).



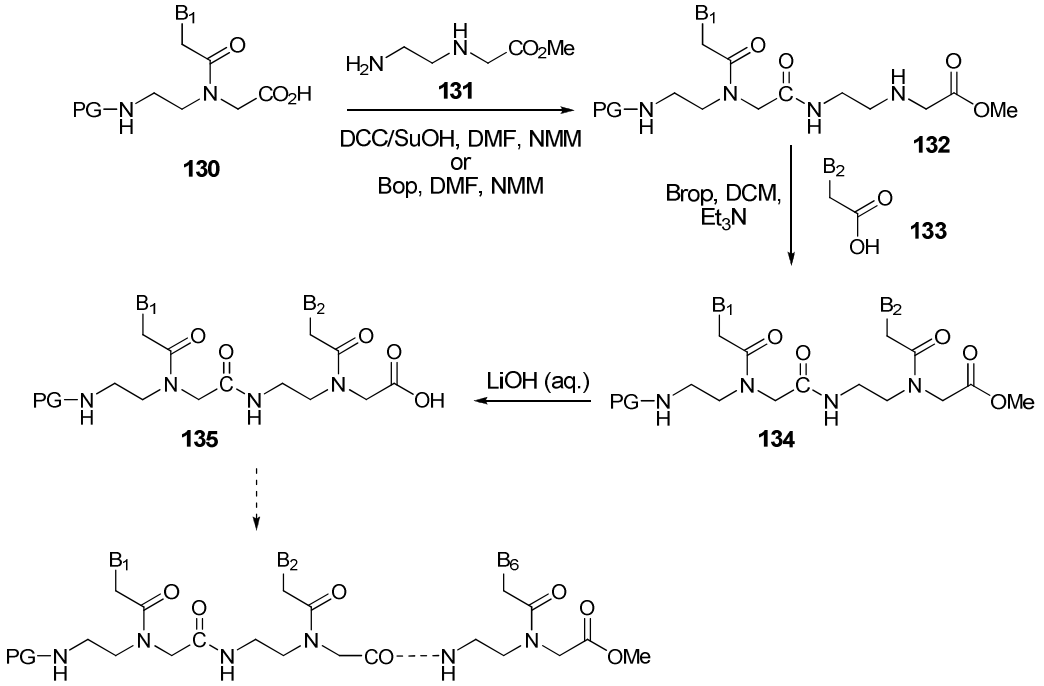
Scheme 39

In order to carry out this coupling step, the two PNA monomers needed to be first deprotected at their *N*- and *C*-terminus respectively. In order to remove the ethyl ester of **84a**, a solution of **84a** in dioxane was treated with a highly diluted aqueous solution of sodium hydroxide (i.e. 0.7 M).[179] After work-up and purification by column chromatography, carboxylic acid **127** was obtained in an 89% yield. In ^{1}H NMR spectrum recorded for **127** the disappearance of the split signals at 1.26 ppm and 1.31 ppm, and 4.18 ppm and 4.26 ppm corresponding to the ethyl ester in **84a** confirmed the cleavage of the ester group. This observation was also evidenced on the ^{13}C NMR spectrum with the disappearance of the split signals at 14.0 ppm and 14.1 ppm, and at 61.7 ppm and 62.3 ppm. The Boc group of the thyminyl PNA monomer **70** was cleaved under acidic conditions, to yield the trifluoroacetic acid salt derivative **82** in a quantitative yield.

Subsequently, **127** and **82** were coupled using HATU as a coupling agent in the presence of DIPEA. After work-up and purification by column chromatography, dimer **128** was obtained in 37 % yield. All the signals corresponding to both starting materials could be observed on both ^{1}H and ^{13}C NMR spectra recorded for **128**. For example, in the ^{1}H NMR the signal at 0.88 ppm corresponding to the CH_3 group of ethyl ester function and the signal at 1.85 ppm, integrating to 6 protons corresponded to the two C(5)-CH_3 groups of the thyminyl rings. However, we did not observe the signal corresponding to the NH proton of the new peptide bond but in the ^{13}C NMR spectrum we reasoned that the signal at 168.6 ppm corresponded to the amide carbonyl group therefore confirming the formation of the peptide bond. Finally, a HRMS further confirmed the formation of **128** through the formation of an $[M+NH_4]^+$ ion peak at m/z of 591.2514; a compound with molecular formula $C_{25}H_{35}N_8O_9$ is theoretically expected to produce an *m/z* of 591.2522.

The low yield obtained for **128** was not a surprise since Condom *et al.*[119] have reported low yields for the coupling of PNA monomer using solution phase protocols. In their work they have overcome this problem by developing a slightly different strategy to prepare PNA hexamer in high yields, using solution-phase synthesis.[119, 180] Instead of coupling directly two PNA monomers, these researchers coupled a PNA monomer unit **130** bearing a free *C*-terminus to the ethyl *N*-(2-aminoethyl)glycinate backbone unit **131** yielding intermediate **132** (*Scheme 40*). The nucleobase acetic acid **133** was then coupled in a separate step, affording dimer **134**. Dimer **132** was then hydrolysed to carbox-

ylic acid **135** and the sequence of step was repeated until the hexamer was formed. Therefore, a similar strategy could be investigated in the future for our purpose.



Scheme 40

Despite the low yield of **128** afforded, we still had it in sufficient quantity to carry out the next step. Thus, **128** was treated with the 'lipid-spacer' intermediate **99** under the same 'Click' reaction conditions that we had used previously for the preparation of the monomer models (*Scheme 39*). After work-up and purification using column chromatography, dimer model **129** was obtained in 46% yield. Due the extremely small amount of material isolated (0.012 g) we were not able to record any NMR spectra for **129**. However, HRMS measurement recorded for **129** showed the molecular ion peak at *m/z* 1014.5977 and theoretically a compound of molecular formula $C_{49}H_{80}N_{11}O_{12}$ is expected to produce an *m/z* of 1014.5982.

Due to the poor yields of the coupling steps and the time restrictions, we did not attempt to prepare the diacetylenic dimer analogues. However, with the incorporation of compound **129** into PDA liposomes, we wanted to show that the inclusion of a larger PNA headgroup in addition to the spacer did not hinder formation and the polymerisation of these liposomes. In addition, we hoped that the inclusion of the PEG-like hydrophilic spacer would help to overcome the precipitation problems encountered by Murray.[2]

CHAPTER 3: LIPOSOME INVESTIGATIONS

3 LIPOSOME INVESTIGATIONS

As defined earlier, liposomes are aggregates of an amphiphilic molecular bilayer. Amphiphilic diacetylenes with polar headgroups and extended hydrocarbon chains form liposomes in aqueous solutions with one set of head-groups exposed at the surface and the other enclosing an interior hydration sphere. The incorporation of suitable receptor elements at the surface of such PDA-liposomes, either by covalent linking to the PDA or by physical embedding within the matrix, can produce sensors in which interactions with specific substrates induce changes in polymer conformation with concomitant colour transitions.[181] In this Chapter, the abilities for all the model compounds (prepared in Chapter 2) to be incorporated into PDA-liposomes will be summarised.

3.1 BACKGROUND

Murray[2] has incorporated thyminyl-PNA monomer and oligomers functionalised with stearic and diacetylenic lipid tail, into PDA-liposomes. His investigations started with the incorporation of 100% of model compounds **22** and **136** (*Figure 73*) and compared these results with the control solution made from 10,12-PCDA (**5** also referred to as matrix lipid) only.

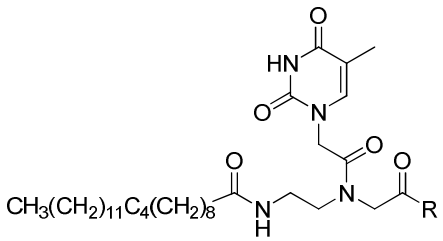


Figure 73

He observed that the acid model **22** incorporated well producing a deep purple solution, while the ester model **136** also resulted in a purple solution, but additionally showed significant precipitation. The electronic spectrum of the polymerised 100% matrix lipid showed visible absorption bands at 639 and 591 nm and a UV absorption band at 241 nm. These absorptions compared well with the absorptions reported by Charych *et al.*,[122] however in their case, absorption below 400 nm was not reported. The electronic spectrum of polymerised acid model **22** showed intense absorption bands at 626, 578,

515 and 273 nm. The band at 273 nm was assigned to the absorption of the thymine residues.[2, 181] The ester model **136** showed the same range of absorption but at lower intensities due to the reduction in concentration due to precipitation. From these results Murray[2] showed that the addition of a PNA monomer head-group did not inhibit or hinder the polymerisation, and that the presence a free carboxylic acid was required in order to prevent precipitation.

Murray[2] reported similar investigations using 1:1 mixtures of the matrix lipid (**5**) and either the acid model **22** or the ester model **136**. Liposomes containing the ester model **136** precipitated upon polymerisation while those containing the acid model **22** gave blue solutions. The electronic spectrum of this blue solution showed absorption bands at 646, 599 and 264 nm. These results demonstrated that model compound **22** was successfully incorporated into PDA-liposomes using **5** as the matrix lipid. Following these results, Murray *et al.*[2, 181] successfully prepared liposomes according to the ratios used by Charych *et al.*,[98] i.e. 95% of matrix lipid and 5 % of sialic-model.

3.2 LIPOSOME PREPARATION

Before presenting the results obtained for the incorporation of the lipid-spacer-PNA monomer models, the protocol for the preparation of liposome is described. It follows the one that has been described by Charych *et al.*.[98] Separate solutions of 10,12-PCDA (**5**) and models were prepared in chloroform. Then, the appropriate amounts of **5** and of the chosen model were measured using micro syringes and were mixed in a glass test tube. The chloroform was subsequently evaporated using a gentle stream of nitrogen gas, to obtain a thin film of lipid at the bottom of the test tube. Consequently, the appropriate amount of deionised water was added so that the total lipid concentration was 1 mM. The sample was then heated to *ca.* 80 ^{o}C using a hot water bath. The titanium tip of the sonicating probe (Sonozap®) was then inserted in the aqueous solution and maintained at *ca.* 1 cm from the bottom of the test tube. Finally, the solution was sonicated for *ca.* 15 to 20 minutes. Charych *et al.*[91] reported that, after sonication, the warm solution should be filtered through a 0.8 μm nylon filter to remove undispersed lipid and traces of titanium particles. Initially this step was also repeated in our case, but after UV irradiation of the filtrate, the solution had only a pale blue shade. It appeared that the filtration had removed all lipids, i.e. undispersed and preformed vesicles. Therefore, the filtration step was avoided in consecutive attempts. Thus, after sonication the test tube

was corked and stored overnight at 4 ^{o}C to cool down. The solution obtained was opaque but uncoloured. After this time, the solution was warmed up to room temperature and purged with nitrogen gas for five minutes. Small aliquots of solutions were then placed in a sample holder and placed at *ca.* 3 cm distance under a UV lamp (254 nm) for up to 30 minutes. After this time, the solutions colour ranged from deep blue to purple (depending on the time of UV exposure). Solutions were subsequently stored in glass vials at 4 ^{o}C for several months.

3.3 INCORPORATION OF LIPID-SPACER-PNA MONOMER MODELS

Therefore, we based our preliminary investigations on the results reported by Murray *et al.*.[181] Although the structures of the models described in Chapter 2 are slightly different from those presented by Murray[2] we believed that the inclusion of the hydrophilic spacer would not hamper the polymerisation process. In fact, Charych *et al.*[98] have included a similar spacer in the receptor molecules used for the detection of the influenza virus. Therefore, the preliminary investigations involving the preparation of liposomes composed entirely of each model compounds was not attempted.

Our studies started with the preparation of liposomes composed of 10,12-PCDA (**5**) only. The results obtained will serve as a control for future incorporation. It was also necessary to familiarise ourselves with the protocols before using our lipid-spacer-PNA monomer models.

Therefore, after 20 minutes of UV irradiation, the opaque liposomes solution obtained from lipid **5** became deep blue in colour (*Figure 74*).

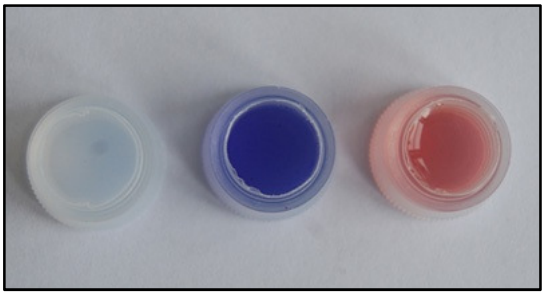


Figure 74 – Liposome solution of 100% 5, before UV irradiation (left), after (right) and after addition of 2 M (aq.) NaOH.

As highlighted in Section 1.4.2, the colour change observed for PDA-liposomes arise due to a conformational change in the PDA backbone which is caused by environmental stimuli, i.e. temperature, pH, mechanical stress. Therefore, on addition of a few drops of a 2 M aqueous solution of sodium hydroxide, the blue liposome solutions made of **5** only immediately changed colour to pink (right vessel *Figure 74*). This colour transition was also observed on the electronic absorption spectra recorded for liposome solutions of **5**, before and after, addition of 2 M aqueous sodium hydroxide solution (*Figure 75*).

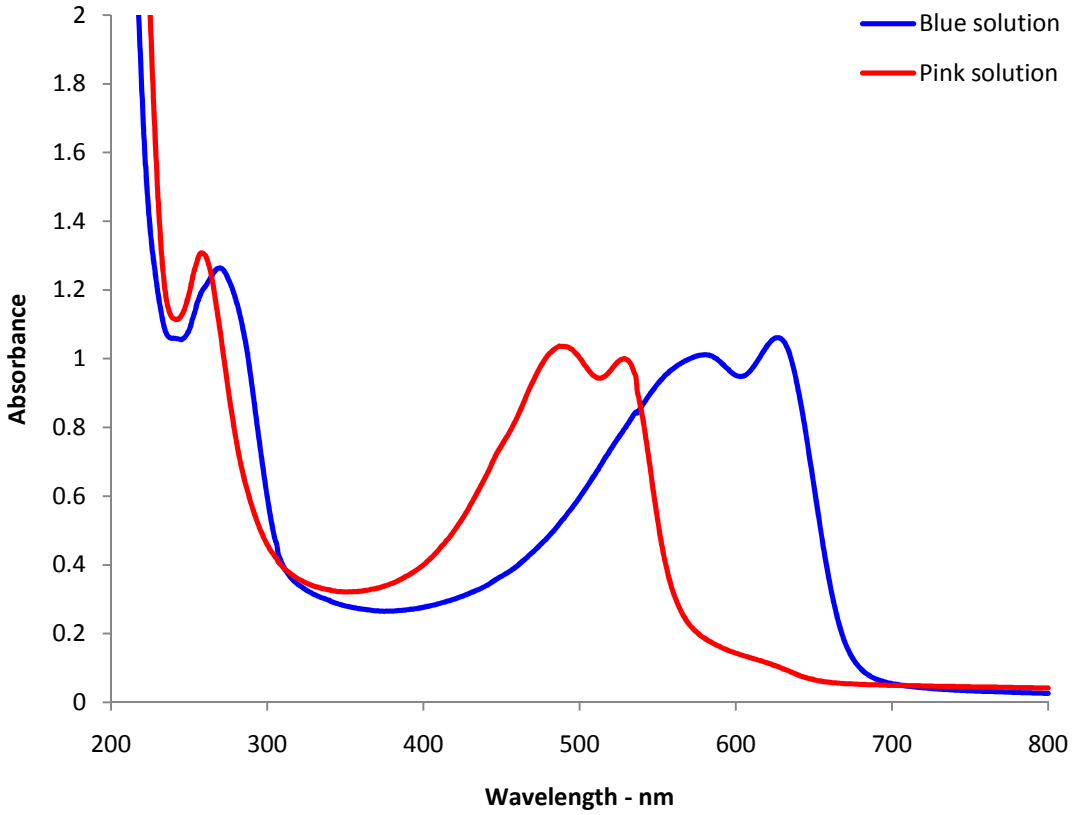
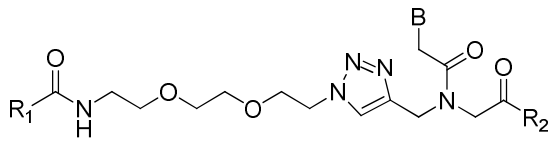


Figure 75 – Electronic absorption spectra of 1 mM liposome solutions of ***5*** *before (blue line) and after (red line) lowering the pH.*

The electronic spectrum of the blue solution showed absorption maxima at 627, 580 and 270 nm, while the red solution absorption maxima were at 529, 487 and 258 nm. These compare well with the electronic absorption spectra reported by Charych *et al.*.[91]

Following this, mixtures composed of 95% matrix lipid **5** and 5% of the thyminyl PNA monomers **95a** and **125a** (*Figure 76*) were prepared separately, and the liposomes were constructed following the protocols described earlier.



R1	B	R2	
		-OCH2CH3	-OH
$CH_3(CH_2)_{11}C_4(CH_2)_8$	T	**95a**	**125a**
	C^{Cbz}	**95b**	**N / A**
	A^{Cbz}	**95c**	**125c**
$CH_3(CH_2)_{16}$	T	**96a**	**126a**
	C^{Cbz}	**96b**	**N / A**
	A^{Cbz}	**96c**	**126c**

Figure 76

After polymerisation, these solutions were blue to purple (depending on the irradiation time) and indistinguishable by eye, from the control liposome made of 100% matrix lipid **5** (*Figure 77*).

Unlike Murray[2] our ester model **95a** incorporated properly into the PDA-liposome, showing similar absorbance to the acid model **125a**, and no precipitation was observed. This indicates that the inclusion of the hydrophilic spacer has improved the hydrophilic/hydrophobic balance of the ester model **95a**. The electronic absorption spectra of these solutions were also very similar to that of the control liposome. The thymine moiety gave rise to a small shoulder on the absorption band in the UV region at *ca.* 270 nm. The resulting absorption bands maxima (λ_{max}) are reported in *Figure 77*.

As mentioned earlier, the incorporation of a sensing element into PDA-liposomes does not necessarily require the element to be covalently linked to the PDA framework. It could simply be physically embedded through hydrophobic interactions. Murray[2] have already studied and reported that PNA monomer model compounds bearing a saturated lipid tail could be incorporated into PDA-liposomes, without inhibiting the polymerisation. Therefore, this work was repeated with our own models **96a** and **126a**. These liposome solutions were visually indistinguishable from the control liposome solution and the electronic absorption spectra showed no difference with the corresponding diacety-

lene analogues (*Figure 77*). Furthermore, no precipitation was observed with the ester model **96a**.

Liposome	λmax (nm)		
poly-**5**	627	580	270
poly-**5** (95%) + poly-**95a** (5%)	623	577	268
poly-**5** (95%) + poly-**125a** (5%)	623	579	275
poly-**5** (95%) + poly-**96a** (5%)	630	586	288
poly-**5** (95%) + poly-**126a** (5%)	615	572	273

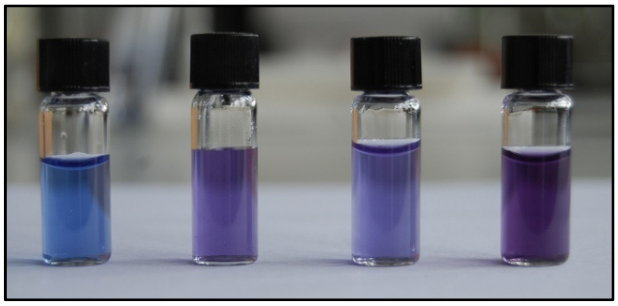


*Figure 77 – Liposome solutions of matrix lipid **5** and 5% monomer models: (left to right) **96a**, **126a**, **95a** and **125a**.*

At this stage, it was necessary to verify if the addition of the novel recognition element would interfere with the colour change. Thus, a few drops of 2 M aqueous sodium hydroxide were added to each sample vial and the colour of the solutions immediately changed to pink or orange (depending on the colour of the starting solution – *Figure 78*). The electronic absorption spectra further evidenced these changes (*Figure 79*).

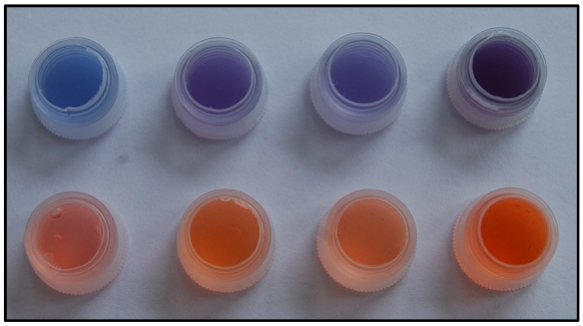


*Figure 78 - Liposome solutions of matrix lipid **5** and 5% monomer models: (left to right) **96a**, **126a**, **95a**, **125a**, (top to bottom) before and after addition of 2 M NaOH.*

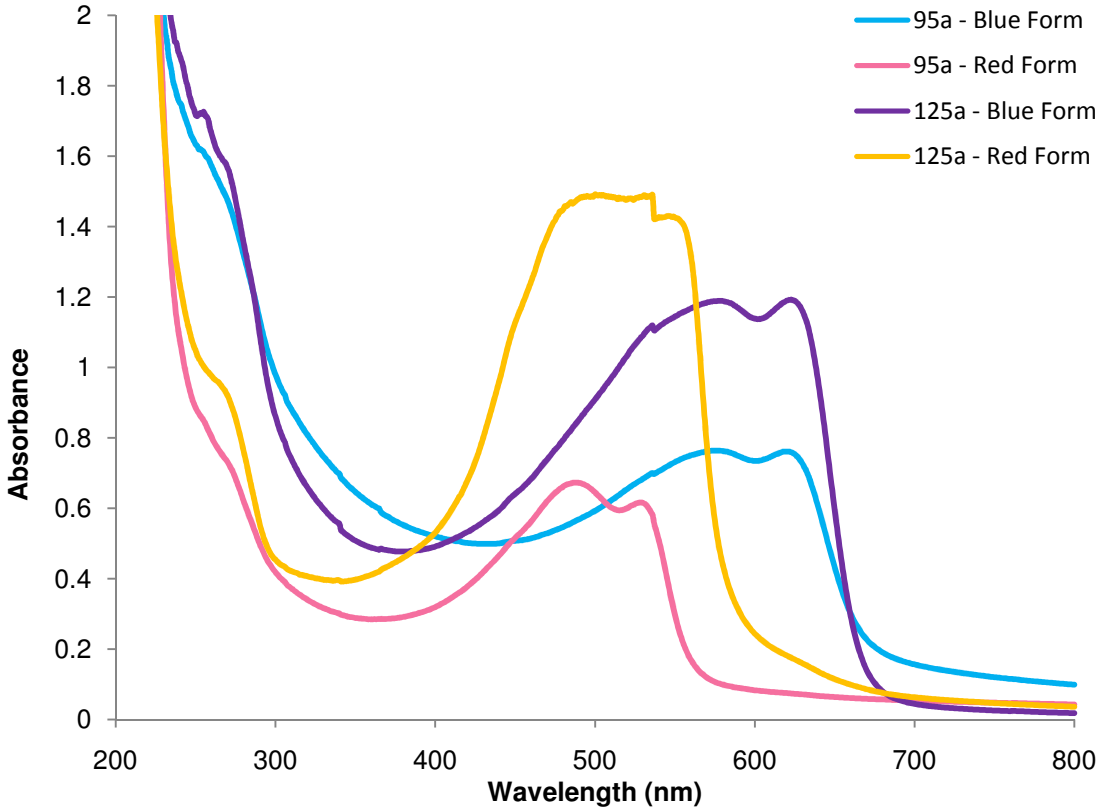
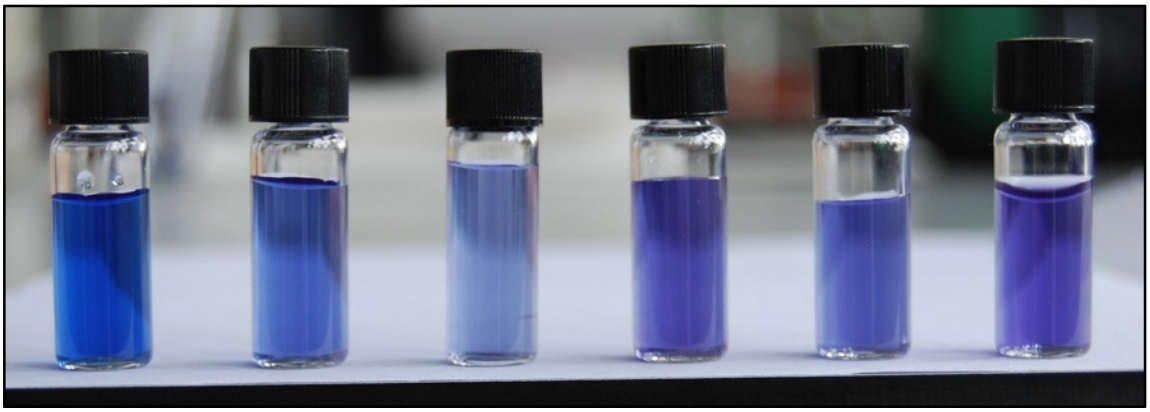


*Figure 79 - Electronic absorption spectra of 1 mM liposome solutions of **95a** and **125a** before (blue and purple) and after (pink and orange) lowering the pH.*

From the electronic spectra in *Figure 79*, it was clear that the purple solution gave rise to higher absorbance than the blue solution. However, the corresponding wavelength values (λ_{max}) were very similar for both the ester and the acid model. Comparable results were obtained for the monomer models bearing the saturated lipid chain (**96a** and **126a**).

The next step in our study involved the incorporation of models bearing cytosinyl or adeninyl PNA monomers. However, since Murray[2] had not undertaken the incorporation of their analogous models, we could not make any comparisons. Therefore, liposomes made of mixtures composed of 95% matrix lipid **5** and 5% of cytosinyl PNA monomers **95b** and **96b** or adeninyl PNA monomers **95c**, **96c**, **125c** or **126c** (*Figure 76*) were prepared separately, following the same protocol as for the thyminyl models. After polymerisation, the solutions colour varied from blue to purple, depending on the irradiation time, but they were indistinguishable by eye from the control liposome made of 100% matrix lipid **5** (*Figure 80*).



Liposome	λmax (nm)		
poly-**5** (95%) + poly-**95b** (5%)	628	584	269
poly-**5** (95%) + poly-**96b** (5%)	634	587	289
poly-**5** (95%) + poly-**95c** (5%)	624	581	257
poly-**5** (95%) + poly-**125c** (5%)	634	582	270
poly-**5** (95%) + poly-**96c** (5%)	627	583	270
poly-**5** (95%) + poly-**126c** (5%)	628	581	270

*Figure 80 - Liposome solutions of matrix lipid 5 and 5% monomer models: (left to right) control **5**, **96b**, **95b**, **96c**, **95c**, **126c** and **125c**.*

On addition of a few drops of 2 M aqueous sodium hydroxide the same colour change was observed. The colour change of PDA-liposomes can also be induced by an increase of the temperature (i.e. thermochromism). Therefore, a blue sample made from 95% of matrix lipid **5** and 5% of model **125c** was equilibrated for ten minutes after each temperature increase (i.e. 30 ^{o}C, 40 ^{o}C, 50 ^{o}C, 60 ^{o}C and 70 ^{o}C) and the UV-vis absorption spectrum of the resulting solutions was recorded. It appeared that although the visible colour change was not readily observed by naked eye (between room temperature and 40 ^{o}C), the variations were more obvious on the electronic absorption spectrum. The results obtained were collected in *Figure 81*.

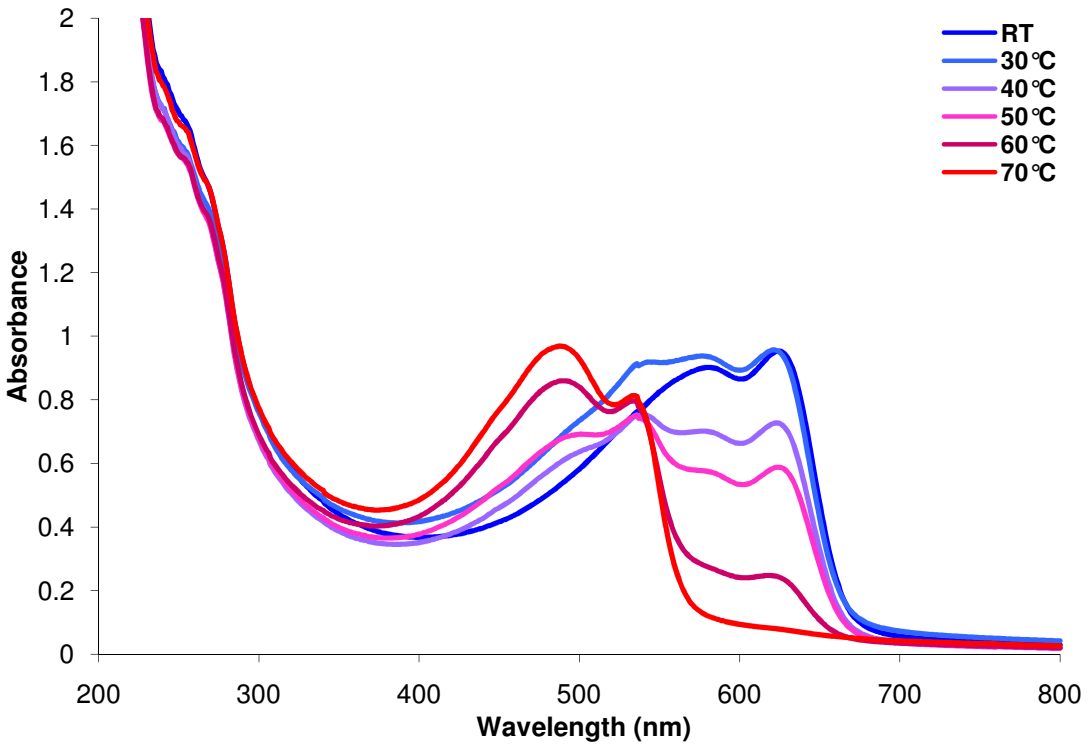


*Figure 81 – Electronic absorption spectra of liposome solutions of matrix lipid **5** and 5% **125c** at different temperatures*

The shift of the wavelength maxima is evident in *Figure 81*. As the temperature increased, the absorbance maxima of the 'blue-form' decreased and the absorbance maxima of the 'red-form' consequently increased. At intermediate temperatures, the PDA-liposomes framework contained a mixture of both the 'blue-' and the 'red-form' with the maxima of each forms still present.

Due to time constraints, we were unable to study the ability of dimer model compound **129** to incorporate into PDA-liposomes. Further benefits of the inclusion of the hydrophilic spacer to overcome the precipitation problems experienced by Murray.[2] would have been demonstrated with this last experiment. If time had not been a problem we would have also been able to study their ability to exhibit the colour transition upon hybridisation to complementary nucleobase using poly(d)A, for example. However, the preliminary studies carried out by Murray were inconclusive and lengthy investigations of the correct conditions were necessary. Even though, the studies using PNA monomer models bearing the hydrophilic spacer proved important in gaining experience in construction of PDA-liposomes. The inclusion of either thymine, cytosine of adenine did not hinder polymerisation or prevent the pH and temperature dependent colour transition of the PDA-liposomes from occurring.

CONCLUSIONS AND FUTURE WORK

4 CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

The combination of preliminary results obtained by Murray[2] and the work described in this research project, has laid foundations for the preparation of polydiacetylene-based nucleic acid biosensors. The principal aim of the project was to improve the hydrophilic/hydrophobic balance of the first generation of PNA-functionalised PDA-liposomes that Murray had synthesized. In fact, they were reported to precipitate from the aqueous solution. From literature precedent by Charych *et al.*[98] and Schneider *et al.*[125] it was reasoned that the inclusion of a PEG-like spacer between the lipid tail and the PNA head-group would overcome precipitation problems endured by Murray.[2] Our first investigations focused on the synthesis of a suitable spacer, and the ethylene glycol derivative, 8-amino-3,6-dioxaoctanoic acid (AEEA - **39**) was our first target.

Compound **39** was chosen because its structure would have allowed peptide connections between its carboxylic acid and the *N*-terminus of the PNA head-group using a peptide bond, as well as between its amino group and the carboxylic acid function of the chosen lipid (i.e. 10,12-PCDA or stearic acid). Thus, our initial efforts were concentrated on the development of an efficient synthetic pathway for the preparation of the Boc-protected AEEA spacer **40**. A number of routes towards **40** were attempted but all afforded limited results. The *O*-alkylation of Boc-protected **46** with ethyl bromoacetate in the presence of sodium hydride or cesium hydroxide only afforded ester **47** in low yields (*ca.* 15%). The same type of reaction of the di-Boc protected **56** again gave only limited amount of the corresponding ether (**57**). Finally, the *O*-alkylation of the di-benzyl analogue **43** afforded **59** in a good yield but this pathway had one limitation, which was the efficient cleavage of the benzyl protecting groups from carboxylic acid analogue **45**. The equipment available at the department at that stage of the project did not allow safe high pressure hydrogenation and therefore, many alternatives were attempted. Due the use of compound **47** in the next steps of the research project, it was mandatory to obtain it in high yields.

With the limited amounts of **47** collected we managed to couple the deprotected amino group of **47** to the lipid tails through amide bonds in reasonable yields (Strategy A). However, a inadequate yield was obtained when di-benzyl carboxylic acid **45** was coupled to *N*-terminus of PNA monomer **70** (Strategy B).

Due to the poor yields for the production of Boc-protected ester **47** and for the peptide coupling according to Strategy B, the group started screening for a different type of connection to be used. Quickly, we became aware of an emerging type of connection, in increasing use in screening chemistry. This reaction is the Huisgen 1,3-dipolar cycloaddition of azides and alkynes which had been revisited and separately reported by Sharpless *et al.*[158] and Meldal *et al.*[159] in 2002. These two research groups have established that the copper(I) catalysed alternative (also called 'click' reaction) allowed the preparation of [1,2,3]-triazole heterocycles both regioselectively (1,4- *versus* 1,5-regioisomer) and in high yields.

With this in mind, our focus moved towards the application of the 'Click' reaction for our purpose. The subsequent design of the novel hydrophilic spacer and the novel PNA monomer were required. It was decided that the azide function would be installed on a Boc-protected PEG-like spacer, analogous to **40**. The alkyne moiety required for the reaction would be at the *N*-terminus of the PNA monomer analogue. Efficient synthetic routes were consequently developed to 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctan-1-azide (**83**) and *N*-alkynyl PNA monomers bearing thymine, Cbz protected cytosine and Cbz protected adenine (**84a-c**). The analogous synthesis exploring the guanine PNA monomer was not attempted as this purine possesses too many synthetic problems. It was deemed that this analogue was not really required at this stage of the initial investigations. With the novel spacer and PNA monomer synthesis nearly optimised, the majority of the following work involved the exploration of the three strategies that have been identified for the preparation of 'lipid-spacer-PNA monomer' models (**95** and **96**).

Our first strategy involved the application of the 'click' reaction to first connect the azide **83** to the alkyne of the PNA monomer (**84a-c**). This intermediate (**97**) was then functionalised with pentacosa-10,12-diynoyl fluoride (**78**), prepared accordingly with Murray.[2] This method was efficient for the preparation of the monomer models bearing thymine, Cbz protected cytosine and Cbz protected adenine (**95a-c**).

Our second strategy implicated the reaction between compound **78** and 8-amino-3,6-dioxaoctan-1-azide hydrochloride (**122**). The azide that was still present in the reaction product **98** was subsequently reacted with the alkyne of the PNA monomers (**84a-c**) using the 'click' reaction conditions, affording models **95a-c**. Compound **98**, however, underwent solid state polymerisation which restricted its use in the next 'Click' reaction

step. This strategy was repeated using stearoyl fluoride (**79**), yielding intermediate **99**. Using this strategy, the saturated lipid-spacer-PNA monomer models **96a-c** were prepared in good yields (above 60%).

Finally, the last strategy led to the preparation of the intermediate ethyl 2-{[1-(8-stearamido-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}aminoacetate trifluoroacetic acid **124**, which contained the backbone unit of a PNA monomer coupled to the spacer through the 1,4-disubstiuted [1,2,3]-triazole. The spacer was also connected to the stearoyl lipid chain through an amide bond. The secondary amine (obtained as a salt) on the backbone section of **122** was available for subsequent attachment of the appropriate nucleobase acetic acid derivatives. This strategy was only partially investigated to afford model **96b** in a reasonable yield (*ca.* 50%). This strategy would have to be further investigated for the preparation of the other lipid-spacer-PNA monomer models. It would be interesting to also verify the photochemical stability of the diacetylenic analogue of **122**.

Having developed models bearing PNA monomers as headgroup, a thyminyl PNA dimer (**128**) was then prepared. Its preparation was carried out by coupling the carboxylic acid of *N*-alkynyl PNA monomer (**127**), to the *N*-terminus of the PNA monomer **82**, using HATU as the activating reagent. The reaction only gave a poor yield (less than 40%). The reason for this low yield has been reported by Condom *et al.*.[180] A different strategy for the preparation of **128** was suggested but not attempted due to time constraints. Nevertheless, the small amount of dimer **128** was sucessfully functionalised using the stearoyl spacer **99** using the 'Click' reaction conditions. This step yielded to the dimer model compounds **129**.

Finally, the last stage of the project focused on the investigation of the preparation of PDA-liposomes using the new monomer model compounds prepared. Both ester (**90a-c** and **91a-c**) and acid models (obtained after basic hydrolysis, **125a/c** and **126a/c**) were successfully incorporated into PDA-liposomes (5% of the model in 95% of matrix lipid **5**). The solutions obtained after photopolymerisation were blue in colour and indistinguishable from the control solution made from 100% 10,12-pentacosadyinoic acid (**5**). No significant differences between the diacetylenic and saturated models in terms of liposome formation could be observed.

These preliminary results demonstrated that the inclusion of the novel spacer **83** did not hinder the formation of liposomes and the polymerisation process. Murray[2] have observed precipitation for their ester monomer analogues. In our case, the absence of precipitation suggested that the hydrophilic spacer had helped to improve the solvation of the 'lipid-spacer-PNA monomer' models. However, in order to fully assess the advantages of including a hydrophilic spacer, it would have been necessary to also incorporate the dimer model **129** into such PDA-liposomes. Unfortunately, due to time and instrumental restrictions the full assessment of the hydrophilic spacer was not completed.

Unfortunately we have not made on time to also incorporate PNA oligomer models, it is still worth noting that hydrophilic spacer **83** alone might not impart enough hydrophilic character. Schneider *et al.*[125] have, indeed, reported the use of two units of 8-amino-3,6-dioxaoctanoic acid and four glutamic acid residues to obtain an amphiphile with enough hydrophilic/hydrophobic balance to prepare liposome. Additionally, the incorporation of such models into liposomes necessitates further investigations. We used the electronic absorption spectra obtained from the liposome solutions, and reproduced in this thesis, to provide conclusions regarding the incorporation of PNA monomer models into the PDA liposome. However, it would be important to find and develop a more conclusive method for examining the composition of the liposome. Differential scanning calorimetry (DSC) and transmission electron microscopy (TEM) are two analytical techniques that have been used by Charych *et al.* to investigate PDA-based liposomes.[68, 69] Finally, in future experiments, it may be interesting to prepare a wide range of lipid-spacer-PNA oligomers models for evaluation as nucleic acid biosensors. For example, it would be interesting to investigate the functionalisation and incorporation into PDA-liposomes of the PNA oligomer H-GGG GCA GTG CCT CAC AA-NH_2, used by Wang *et al.*[7] for the detection of a common point mutation in the p53 gene.

EXPERIMENTAL

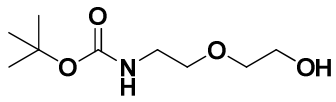
5 GENERAL EXPERIMENTAL

Commercially available reagents from Sigma-Aldrich, Alfa Aesar and NovaBiochem companies were generally used as supplied without further purification. Unless otherwise stated, all solvents were used without further purification. Solvents were dried following standard procedures.[182] Light petroleum refers to the fraction boiling between 40 ºC and 60 ºC. Reactions were routinely carried out under an inert atmosphere of argon or nitrogen. Analytical thin layer chromatography (TLC) was performed on Merck aluminium backed plates coated with Merck Kieselgel 60 GF_{254} (Art. 05554). Developed plates were visualised under ultra-violet light (254 nm) and/or with an alkaline potassium permanganate dip. Flash column chromatography was performed using DAVISIL® silica (60 Å; 35-70 μm) purchased from Fisher (cat. S/0693/60). Fully characterised compounds were chromatographically homogeneous.

Melting points were determined using a Stuart Melting Point SMP10 apparatus and are uncorrected. IR spectra were recorded on a Perkin-Elmer 1600 FT IR spectrometer. Spectra were recorded as potassium bromide discs or as films between sodium chloride plates. Bands are defined as br (broad), s (strong), m (medium) and w (weak). UV-vis spectra were recorded on a Shimadzu UV-160 spectrophotometer. NMR spectra were recorded on a Bruker AC200 at 200 MHz (^{1}H) and 50 MHz (^{13}C), or a Bruker DPX400 at 400 MHz (^{1}H) and 101 MHz (^{13}C). Chemical shifts are reported in parts per million (δ in ppm) relative to tetramethylsilane (0 ppm) and are referenced against solvent residual resonances (δ_H 7.26 for chloroform, δ_H 2.50 for dimethyl sulfoxide, δ_H 4.79 deuterium oxide and δ_H 3.31 for methanol) for ^{1}H NMR spectra and solvent signals (δ_C 77.16 for chloroform, δ_C 39.52 for methyl sulfoxide and δ_C 49.00 for methanol) for ^{13}C NMR spectra.[183] Chemical shift values are accurate to ±0.01 ppm and ±0.1 ppm respectively. Coupling constants (*J*) are reported in Hertz (Hz). Multiplicities are described as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), dt (doublet of triplet) and br (broad signal). In ^{13}C NMR spectra, signals corresponding to CH, CH_2 or CH_3 groups are assigned from DEPT. Products containing amide bonds were isolated as mixtures of two rotamers, several NMR signals of these products were doubled. Low resolution (LRMS) and high resolution mass spectra (HRMS) were obtained using electron ionisation (EI) or electrospray (ESI) and were performed at the EPSRC National Mass Spectrometry Service Centre, University of Wales, Swansea. Elemental analyses were car-

ried out by the analytical service within the Chemistry Department at Heriot-Watt University using an Exeter CE-440 Elemental Analyser.

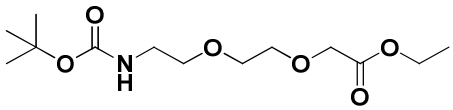
Preparation of 5-(*tert*-butoxycarbonyl)amino-3-oxapentan-1-ol (46)



$C_9H_{19}NO_4$

A solution of 5-amino-3-oxapentan-1-ol (**42**; 6.29 g, 60.0 mmol) in water (200 mL) was cooled in an ice bath. A solution of di-*tert*-butyl dicarbonate (13.1 g, 60.1 mmol) in dioxane (3 mL) was added and the temperature was allowed to rise to room temperature whilst the pH was maintained at 10.5 by the addition of 2 M sodium hydroxide solution. After 2 h, the mixture was concentrated to a paste and triturated with dichloromethane (300 mL). The suspension was filtered and the organic phase was dried with anhydrous magnesium sulphate. After filtration, the solvent was removed under reduced pressure to afford the title compound (**46**; 13.7 g, 66.7 mmol; 90%) as a clear oil; R_f (dichloromethane-methanol 95:5) = 0.28; υ_{max}(film)/cm^{-1} 3364br s, 2977s, 2934s, 2875s, 1689s, 1529s; δ_H (200 MHz, $CDCl_3$) 1.34 (9H, s, (C*H*$_3$)$_3$), 2.98 (1H, br s, O*H*), 3.25 (2H, dt, $^3J_{H\text{-}NH}$ 4.9 and $^3J_{H\text{-}CH2}$ 5.4, NHC*H*$_2$CH$_2$), 3.43-3.49 (4H, m, C*H*$_2$OC*H*$_2$), 3.60-3.63 (2H, m, C*H*$_2$OH), 5.45 (1H, br t, N*H*CH$_2$); δ_C (50 MHz, $CDCl_3$) 28.3 ((*C*H$_3$)$_3$), 40.2 (*C*H$_2$), 61.1 (*C*H$_2$), 70.2 (*C*H$_2$), 70.2 (*C*H$_2$), 79.2 (*C*(CH$_3$)$_3$), 155.1 (*C*ONH); *m/z* (ESI) 206 ([M+H]$^+$, 40%), 167 (100), 150 (44), 106 (28), 44 (80); HRMS (ESI) (Found: [M+H]$^+$, 206.1390. $C_9H_{20}NO_4$ requires *m/z*, 206.1387).

Preparation of 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctanoic acid ethyl ester (47)



$C_{13}H_{25}NO_6$

Method A

A solution of 5-(*tert*-butoxycarbonyl)amino-3-oxapentan-1-ol (**46**; 1.49 g, 7.26 mmol) in anhydrous tetrahydrofuran (15 mL) was slowly added to a stirred mixture of NaH (60% in mineral oil, 0.32 g, 8.00 mmol) in anhydrous tetrahydrofuran (30 mL). The mixture was heated at reflux for 75 min and then cooled to between -5 to -10 °C whereupon ethyl bromoacetate (1.3 g, 8.0 mmol) was slowly added. The reaction mixture was then stirred overnight at room temperature. After this time the reaction was quenched by addition of water (10 mL) and tetrahydrofuran was concentrated under reduce pressure. The aqueous solution was extracted with ethyl acetate (5 × 20 mL) and the combined organic layers were washed with water (2 × 10 mL). The aqueous layers were back-extracted with ethyl acetate (2 × 20 mL) and the combined organic extracts were washed with brine (10 mL) before being dried over anhydrous magnesium sulphate and filtrated. The solvent was removed in vacuo to give a crude, dark yellow liquid which was subjected to flash column chromatography (ethyl acetate-light petroleum ether 1:1 to 9:1) to afford the title compound (**47**; 0.21 g, 0.72 mmol; 9.9%) as a yellow liquid.

Method B

The same procedure was used as method A except that **46** (4.31 g, 21.0 mmol) was reacted with NaH (60% in mineral oil, 1.02 g, 25.2 mmol) in anhydrous tetrahydrofuran (50 mL) at room temperature for 15 min, before the slow addition of ethyl bromoacetate (4.21 g, 25.2 mmol). The work-up procedure was as described in method A. The title compound (**47**; 0.58 g, 1.99 mmol; 10%) was obtained as a yellow liquid.

Method C

The same procedure was used as method A except that **46** (1.74 g, 8.48 mmol) was reacted with NaH (60% in mineral oil, 0.41 g, 10.3 mmol) in anhydrous dimethylformamide (30 mL) at room temperature for 15 min, before the slow addition of ethyl bro-

moacetate (1.84 g, 11.0 mmol). The work-up procedure was as described in method A. The title compound (**47**; 0.46 g, 1.58 mmol; 19%) was obtained as a yellow liquid

Method D

To a suspension containing activated powdered 4 Å molecular sieves (1.00 g) in anhydrous dimethylformamide (15 mL) was added cesium hydroxide monohydrate (0.40 g, 2.38 mmol) and the mixture was stirred under argon for 30 min. A solution of **46** (0.49 g, 2.38 mmol) solution in anhydrous dimethylformamide (10 mL) was then added and the following mixture was stirred for 1 h. After the slow addition of ethyl bromoacetate (0.47 g, 2.40 mmol), the reaction was allowed to proceed at room temperature overnight. After this time dimethylformamide was evaporated under reduced pressure and the paste obtained was coevaporated with toluene (2 × 10 mL) and diethyl ether (2 × 10 mL). Unfortunately, none of the products collected after purification corresponded to **47**.

Method E

To a similar suspension as method D [activated powdered 4 Å molecular sieves (2.00 g) in anhydrous dimethylformamide (15 mL) and cesium hydroxide monohydrate (0.51 g, 3.06 mmol)] was added **46** (0.52 g, 2.53 mmol) and the solution was stirred for 1 h. To this solution was then added ethyl bromoacetate (0.51 g, 3.06 mmol) and tetrabutyl ammonium iodide (1.12 g, 3.03 mmol) and the reaction was allowed to proceed at room temperature for 24 h. After this time the work-up was as described method F and the title compound (**47**; 0.13 g, 0.44 mmol; 18%) was obtained as a yellow liquid.

Method F

To a suspension containing activated powdered 4 Å molecular sieves (1.00 g) in anhydrous acetonitrile (15 mL) was added cesium hydroxide monohydrate (0.52 g, 3.08 mmol) and the mixture was stirred under argon for 30 min. **46** (0.53 g, 2.56 mmol) in anhydrous acetonitrile (10 mL) was then added and the following mixture was stirred for 1 h. After this time ethyl bromoacetate (0.51 g, 3.07 mmol) and tetrabutyl ammonium iodide (1.14 g, 3.07 mmol) were then added and the reaction was allowed to proceed at room temperature for 24 h. After this time the work-up was as described for method F and the title compound (**47**; 0.089 g, 0.31 mmol; 10%) was obtained as a yellow liquid.

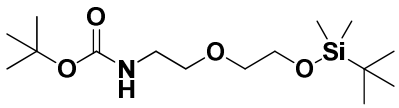
Method G

The same procedure as method F was used except that anhydrous cesium carbonate (1.67 g, 5.12 mmol) was used [**46** (0.53 g, 2.56 mmol) in anhydrous dimethylformamide (10 mL)]. Ethyl bromoacetate (0.51 g, 3.07 mmol) was then added and the reaction was allowed to proceed at room temperature for 24 h. After a similar work-up as method F, none of the products collected after purification corresponded to **47**.

Method H

46 (2.05 g, 9.92 mmol) and anhydrous potassium carbonate (2.80 g, 20.2 mmol) were dissolved in anhydrous dimethylformamide (20 mL) and the heterogeneous solution was stirred for 3 h before the slow addition of ethyl bromoacetate (4.14 g, 24.8 mmol). The reaction mixture was stirred overnight and the work-up was as described in method A. The title compound (**47**; 0.49 g, 16.8 mmol; 17%) was obtained as a yellow liquid; R_f (ethyl acetate-light petroleum 8:2) = 0.30; υ_{max}(film)/cm^{-1} 2981s, 2933s, 2878s, 1749s, 1706s, 1507s; δ_H (200 MHz; $CDCl_3$) 0.85 (3H, t, *J* 7.1, C*H*$_3$), 1.10 (9H, s, (C*H*$_3$)$_3$), 2.88 (2H, br dt, NHC*H*$_2$), 3.11 (2H, t, *J* 5.4, OCH$_2$C*H*$_2$O), 3.23-3.24 (4H, m, C*H*$_2$OC*H*$_2$), 3.71 (2H, s, OC*H*$_2$COO), 3.78 (2H, q, *J* 7.1,OC*H*$_2$), 4.81 (1H, br t, N*H*CH$_2$); δ_C (50 MHz; $CDCl_3$) 14.1 (*C*H$_3$), 28.3 (C(*C*H$_3$)$_3$), 40.2 (NH*C*H$_2$), 60.2 (O*C*H$_2$CH$_3$), 68.5 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.5 (O*C*H$_2$), 70.7 (O*C*H$_2$), 78.8 (*C*(CH$_3$)$_3$), 155.9 (*C*ONH), 170.3 (*C*OO); *m/z* (ESI) 292 ([M+H]$^+$, 88%), 314 ([M+NH$_4$]$^+$, 100), 236 (44), 192 (41), 134 (10); HRMS (ESI) (Found: [M+H]$^+$, 292.1754. $C_{13}H_{26}NO_6$ requires *m/z*, 292.1755).

Preparation of (*tert*-butoxycarbonyl)amino(*tert*-butyldimethylsilyloxy)-3-oxapentan (54)

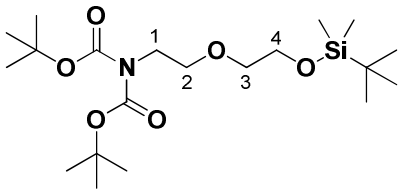


$C_{15}H_{33}NO_4Si$

To a solution of 5-(*tert*-butoxycarbonyl)amino-3-oxapentan-1-ol (**46**; 1.03 g, 5.00 mmol) in anhydrous dimethylformamide (20 mL) was added *tert*-butylchloro-dimethylsilane (0.96 g, 6.4 mmol) and imidazole (0.51 g, 7.5 mmol). The mixture was stirred for 18 h at room temperature. After this time the yellow solution obtained was poured in water and extracted with ethyl acetate (6 × 10 mL). The combined organic

phase was dried over anhydrous magnesium sulphate. After filtration, the solvent was removed to give a colourless oil that solidified to a white solid, which was subjected to flash column chromatography (ethyl acetate-light petroleum ether 20:80) to yield the title compound (**54**; 1.34 g, 4.2 mmol; 84%) as a white solid: m.p. 41-42 °C; R_f (ethyl acetate-light petroleum 2:8) = 0.28; υ_{max}(film)/cm^{-1} 2932m, 2861m, 1715s, 1510s; δ_H (200 MHz; $CDCl_3$) 0.05 (6H, s, Si(C*H*$_3$)$_2$), 0.87 (9H, s, SiC(C*H*$_3$)$_3$), 1.41 (9H, s, OC(C*H*$_3$)$_3$), 3.24-3.28 (2H, m, NHC*H*$_2$), 3.47-3.54 (4H, m, OC*H*$_2$C*H*$_2$O), 3.73 (2H, t, *J* 3.4, NHCH$_2$C*H*$_2$O), 4.98 (1H, br t, N*H*CH$_2$); δ_C (50 MHz; $CDCl_3$) 18.4 (Si(*C*H$_3$)$_2$), 25.9 (SiC(*C*H$_3$)$_3$), 28.4 (OC(*C*H$_3$)$_3$), 40.4 (NH*C*H$_2$), 62.7 (O*C*H$_2$), 70.1 (O*C*H$_2$), 72.3 (O*C*H$_2$), 79.1 (O*C*(CH$_3$)$_3$ and Si*C*(CH$_3$)$_3$), 155.9 (*C*ONH); *m/z* (ESI) 320 ([M+H]$^+$, 54%), 337 ([M+NH$_4$]$^+$, 100), 264 (41), 220 (15); HRMS (ESI) (Found: [M+H]$^+$, 320.2251. $C_{15}H_{33}NO_4Si$ requires *m/z*, 320.2252).

Preparation *N,N*-bis(*tert*-butoxycarbonyl)amino(*tert*-butyldimethylsilyloxy)-3-oxapentan (55)

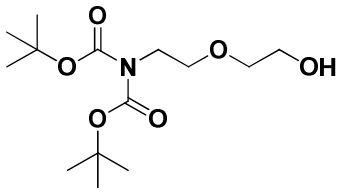


$C_{20}H_{41}NO_6Si$

Compound **54** (0.12 g, 0.38 mmol), 4-dimethylaminopyridine (0.046 g, 0.38 mmol) and di-*tert*-butyl dicarbonate (0.13 g, 0.58 mmol) were dissolved in anhydrous acetonitrile (5 mL) leading to a pale yellow solution. The reaction mixture was stirred at room temperature for 24 h, after which time the solution was dark yellow and TLC analysis (ethyl acetate-light petroleum 1:9) showed that the reaction was not completed, therefore another portion of di-*tert*-butyl dicarbonate (0.124 g, 0.57 mmol) was added and the reaction mixture was stirred for another 24 h. After this time, the solution was deep orange, it was poured on water and extracted with ethyl acetate (4 × 10 mL). The combined organic phase was successively washed with 1 M hydrochloric acid solution (10 mL) and brine (4 × 10 mL). The combined organic phase was finally dried with anhydrous magnesium sulphate, filtered and the solvent was evaporated. The crude residue was subjected to flash column chromatography (light petroleum ether-ethyl acetate 80:20) to

yield the title compound (**55**; 0.128 g, 0.30 mmol; 80%) as a pale yellow oil; R_f (ethyl acetate-light petroleum 1:9) = 0.34; υ_{max}(film)/cm^{-1} 2979s, 2957s, 2930s, 2860s, 1750s, 1699s; δ_H (400 MHz; $CDCl_3$) 0.21 (6H, s, Si(C*H*$_3$)$_2$), 0.88 (9H, s, SiC(C*H*$_3$)$_3$), 1.48 (18H, s, 2 × C(C*H*$_3$)$_3$), 3.54 (2H, t, *J* 5.5, C*H*$_2$(3)), 3.62 (2H, t, *J* 6.3, C*H*$_2$(2)), 3.75 (2H, t, *J* 5.5, C*H*$_2$(4)), 3.80 (2H, t, *J* 6.3, C*H*$_2$(1)); δ_C (100 MHz; $CDCl_3$) -5.30 (Si(*C*H$_3$)$_2$), 25.9 (SiC(*C*H$_3$)$_3$), 28.0 (C(*C*H$_3$)$_3$), 45.3 (N*C*H$_2$), 62.6 (*C*H$_2$Si), 69.3 (*C*H$_2$(2)O), 72.3 (O*C*H$_2$(3)), 82.2 (*C*(CH$_3$)$_3$ and Si*C*(CH$_3$)$_3$), 152.6 (*C*ONCH$_2$); *m/z* (ESI) 420 ([M+H]$^+$, 25%), 364 (100), 308 (37), 220 (10), 60 (45); HRMS (ESI) (Found: [M+H]$^+$, 420.2776. $C_{20}H_{41}NO_6Si$ requires *m/z*, 420.2781). The identity of the product was deduced from the NOESY spectrum.

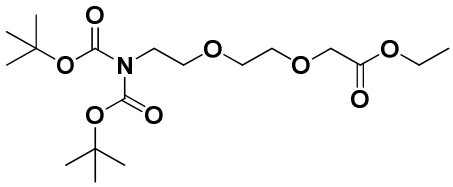
Preparation *N*,*N*-bis(*tert*-butoxycarbonyl)amino-3-oxapentan-1-ol (56)



$C_{14}H_{27}NO_6$

55 (1.05 g, 2.50 mmol) was dissolved in anhydrous tetrahydrofuran (11 mL) and TBAF (5 mL; 1 M in tetrahydrofuran, dried over 4 Å molecular sieves) was added *via* a syringe. The reaction mixture was stirred for 5 h and after this time it was concentrated *in vacuo*. The crude residue was subjected to flash column chromatography (ethyl acetate-light petroleum 7:3) to yield the title compound (**56**; 0.763 g, 2.50 mmol; 100%) as colourless oil; R_f (ethyl acetate-light petroleum 6:4) = 0.34; υ_{max}(film)/cm^{-1} 3321br s, 2979s, 2956s, 2928s, 2861s, 1762s, 1689m; δ_H (400 MHz; $CDCl_3$) 1.49 (18H, s, 2 × C(C*H*$_3$)$_3$), 2.47 (1H, br s, O*H*), 3.50 (2H, br dt, C*H*$_2$OH), 3.53-3.62 (4H, m, NC*H*$_2$CH$_2$OC*H*$_2$), 3.72 (2H, t, *J* 5.8 , NCH$_2$C*H*$_2$); δ_C (100 MHz; $CDCl_3$) 27.9 (C(*C*H$_3$)$_3$), 45.4 (N*C*H$_2$), 61.6 (*C*H$_2$OH), 69.3 (O*C*H$_2$), 72.1 (O*C*H$_2$), 82.4 (*C*(CH$_3$)$_3$), 152.9 (*C*OO); *m/z* (EI) 306 (M$^+$, 15%).

Preparation 8-bis(*tert*-butoxycarbonyl)amino-3,6-dioxaoctanoic ethyl ester (57)



$C_{18}H_{33}NO_8$

Method A

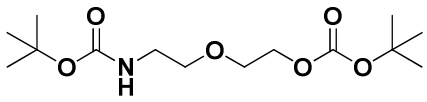
To a slurry of sodium hydride (60% in mineral oil; 0.017 g, 0.43 mmol) in anhydrous tetrahydrofuran (3 mL) was added a solution of **56** (0.103 g, 0.340 mmol) in anhydrous tetrahydrofuran (3 mL). The mixture was heated to a gentle reflux for 75 min, after being cooled to 0 °C using a ice/salt bath. Ethyl bromoacetate (0.060 mL, 0.51 mmol) was then slowly added to the solution and the stirring was continued for 17 h at room temperature. After this time, water (4 mL) was very carefully added and tetrahydrofuran was removed. The aqueous phase was extracted with ethyl acetate (5 × 10 mL). The combined organic layer was washed with water (2 × 10 mL) and brine (10 mL). The aqueous phase was back-extracted with ethyl acetate (10 mL). The combined organic phase was finally dried over anhydrous magnesium sulphate and the evaporation of the solvent yielded a pale yellow oil. The crude mixture was subjected to flash column chromatography (ethyl acetate-light petroleum 4:6) but ^{1}H NMR spectra of fractions obtained failed showing signals corresponding to **57**.

Method B

56 (0.402 g, 1.32 mmol) was dissolved in anhydrous tetrahydrofuran (5 mL) and ethyl bromoacetate (0.440 mL, 3.90 mmol) was then added. The solution was cooled in an ice bath and sodium hydride (60% in mineral oil; 0.0790 g, 1.98 mmol) was carefully added in small portions. The reaction was stirred for 17 h while it was allowed to slowly reach room temperature. After this time, water (5 mL) was very carefully added and tetrahydrofuran was removed. The aqueous phase was extracted with ethyl acetate (5 × 10 mL). The combined organic layer was washed with water (2 × 10 mL) and brine (10 mL). The combined organic phase was finally dried over anhydrous magnesium sulphate and the evaporation of the solvent yielded a deep yellow oil. The crude mixture was subjected to flash column chromatography (ethyl acetate-light petroleum 4:6), affording the title compound (**57**; 0.099 g, 0.25 mmol; 19%) as a pale yellow oil; R_f (ethyl

acetate-light petroleum 4:6) = 0.28; δ_H (400 MHz; $CDCl_3$) 1.23 (3H, t, *J* 7.1, C*H*$_3$), 1.45 (18H, s, 2 × C(C*H*$_3$)$_3$), 3.57 (2H, t, *J* 5.8, NC*H*$_2$CH$_2$), 3.59-3.68 (4H, m, OC*H*$_2$C*H*$_2$O), 3.75 (2H, t, *J* 5.8, NCH$_2$C*H*$_2$), 4.09 (2H, s, OC*H*$_2$COO), 4.16 (2H, q, *J* 7.1, OC*H*$_2$); δ_C (100 MHz; $CDCl_3$) 14.8 (*C*H$_3$), 27.9 (C(*C*H$_3$)$_3$), 45.4 (N*C*H$_2$), 61.6 (O*C*H$_2$CH$_3$), 68.7 (O*C*H$_2$), 69.2 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.5 (O*C*H$_2$), 82.4 (*C*(CH$_3$)$_3$), 152.9 (N*C*OO), 169.2 (*C*OO).

Preparation of 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctanoic *tert*-butyl ester (58)



$C_{14}H_{27}NO_6$

Method A

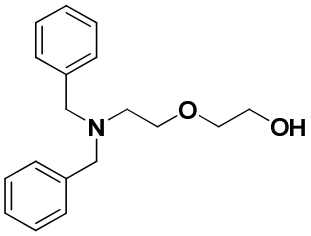
To a slurry of sodium hydride (60% in mineral oil; 0.017 g, 0.43 mmol) in anhydrous tetrahydrofuran (3 mL) was added a solution of **56** (0.103 g, 0.340 mmol) in anhydrous tetrahydrofuran (3 mL). The mixture was heated to a gentle reflux for 75 min, after being cooled to 0 °C using an ice/salt bath. Ethyl bromoacetate (0.060 mL, 0.51 mmol) was then slowly added to the solution and the stirring was continued for 17 h at room temperature. After this time, water (4 mL) was very carefully added and tetrahydrofuran was removed under reduced pressure. The aqueous phase was extracted with ethyl acetate (5 × 10 mL). The combined organic layer was washed with water (2 × 10 mL) and brine (10 mL). The aqueous phase was back-extracted with ethyl acetate (10 mL). The combined organic phase was finally dried over anhydrous magnesium sulphate and the evaporation of the solvent yielded a pale yellow oil. The crude mixture was subjected to flash column chromatography (ethyl acetate-light petroleum 4:6), affording the title compound (**58**; 0.043 g, 0.14 mmol; 40%) as a colourless oil.

Method B

To a suspension of powdered 4 Å molecular sieves (0.18 g) in anhydrous dimethylformamide (6 mL) was added cesium hydroxide monohydrate (0.251 g, 1.49 mmol). This heterogeneous mixture was stirred for 15 min, then *N*,*N*-bis(*tert*-butoxycarbonyl)amino-3-oxapentan-1-ol (**56**; 0.145 g, 0.48 mmol) in anhydrous dimethylformamide (2 mL) was added and the reaction mixture was stirred at room temperature for 1 h. Ethyl bromoacetate (0.265 mL, 2.38 mmol) was then added and the reaction was stirred for 24 h.

After this time, water (6 mL) was added and dimethylformamide was removed under reduced pressure. The residue was dissolved in ethyl acetate (10 mL) and filtered. The organic phase was separated and the aqueous phase was extracted with ethyl acetate (4 × 10 mL). The combined organic layer was washed with water (2 × 10 mL) and brine (10 mL) and finally dried over anhydrous magnesium sulphate. The solvent was evaporated and the crude mixture obtained was subjected to flash column chromatography (ethyl acetate-light petroleum 4:6), affording the title compound (**58**; 0.13 g, 0.43 mmol; 90%) as a pale yellow oil; R_f (ethyl acetate-light petroleum 4:6) = 0.31; δ_H (200 MHz; $CDCl_3$) 1.38 (9H, s, C(C*H*$_3$)$_3$), 1.42 (9H, s, C(C*H*$_3$)$_3$), 3.50 (2H, br dt, NC*H*$_2$CH$_2$), 3.72 (2H, t, *J* 3.4, OC*H*$_2$CH$_2$OCO), 3.58-3.63 (2H, m, NCH$_2$C*H*$_2$), 4.13-4.18 (2H, m, C*H*$_2$OCO), 4.88 (1H, br s, N*H*).

Preparation of *N*,*N*-dibenzyl-5-amino-3-oxapentan-1-ol (43)[126]

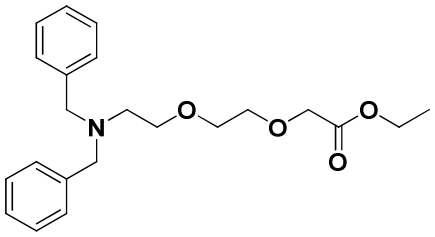


$C_{18}H_{23}NO_2$

5-Amino-3-oxapentan-1-ol (**42**; 3.15 g, 29.9 mmol) and potassium carbonate (8.30 g, 60.1 mmol) were mixed in water (8 mL). Benzyl bromide (7.2 mL, 59.9 mmol) was then added dropwise over 50 min and, after this time, the reaction mixture was stirred at room temperature for 50 h. After this time, diethyl ether (50 mL) was added and the layers were separated. The aqueous layer was extracted with diethyl ether (2 × 20 mL) and the combined organic phase was washed with brine (20 mL) and dried over anhydrous magnesium sulphate. The pale yellow oil obtained on evaporation of the solvent was subjected to flash column chromatography (ethyl acetate-light petroleum 4:6) and the title compound (**43**; 6.82 g, 23.9 mmol; 80%) was obtained as a pale yellow oil; R_f (ethyl acetate-light petroleum 3:7) = 0.28; υ_{max}(film)/cm^{-1} 3421m, 3027m, 2925m, 2871m, 1494s, 1433s; δ_H (200 MHz; $CDCl_3$) 2.76 (2H, t, *J* 5.9, NC*H*$_2$CH$_2$), 2.97 (1H, br s, O*H*), 3.51-3.54 (2H, m, CH$_2$C*H*$_2$OH), 3.61-3.65 (4H, m, C*H*$_2$OC*H*$_2$), 3.72 (4H, s, 2 × NC*H*$_2$Ar), 7.26-7.46 (10H, m, 10 × Ar-*H*); δ_C (50 MHz; $CDCl_3$) 52.6 (N*C*H$_2$), 58.7 (N*C*H$_2$Ar), 61.5 (*C*H$_2$OH), 69.3 (O*C*H$_2$), 71.8 (O*C*H$_2$), 126.7 (Ar-*C*H), 127.9 (Ar-*C*H),

128.6 (Ar-*C*H), 139.1 (Ar-*C*); *m/z* (ESI) 286 ([M+H]$^+$, 100%); HRMS (ESI) (Found: [M+H]$^+$, 286.1802. $C_{18}H_{23}NO_2$ requires *m/z*, 286.1802).

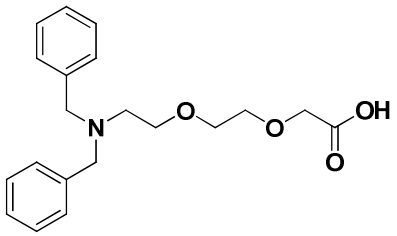
Preparation of *N*,*N*-dibenzyl-8-amino-3,6-dioxaoctanoic acid ethyl ester (59)



$C_{22}H_{29}NO_4$

To a slurry of sodium hydride (60% in mineral oil; 0.48 g, 12 mmol) in anhydrous tetrahydrofuran (15 mL) was added a solution of *N*,*N*-dibenzyl-5-amino-3-oxapentan-1-ol (**43**; 2.86 g, 10.0 mmol) in anhydrous tetrahydrofuran (8 mL). The mixture was heated to reflux for 2 h after being cooled to 0 °C using an ice/salt bath. Ethyl bromoacetate (1.22 mL, 11.0 mmol) was then slowly added to the solution and the stirring was continued for 17 h at room temperature. After this time, water (5 mL) was carefully added and tetrahydrofuran was removed under reduced pressure. The aqueous phase was extracted with ethyl acetate (5 × 20 mL). The combined organic layer was washed with water (2 × 10 mL) and brine (10 mL). The aqueous phase was backextracted with ethyl acetate (20 mL). The combined organic phase was finally dried over anhydrous magnesium sulphate and the evaporation of the solvent yielded a brown liquid. The crude mixture was subjected to flash column chromatography (ethyl acetate-light petroleum 4:6), affording the title compound (**59**; 2.16 g, 5.82 mmol; 40%) as a pale yellow oil; R_f (ethyl acetate-light petroleum 4:6) = 0.38; υ_{max}(film)/cm^{-1} 3027w, 2872m, 1754s, 1494m; δ_H (200 MHz; $CDCl_3$) 1.16 (3H, t, *J* 7.1, C*H*$_3$), 3.61 (2H, t, *J* 6.2, NC*H*$_2$), 3.46-3.52 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.57 (4H, s, 2 × NC*H*$_2$Ar), 4.03 (2H, s, OC*H*$_2$COO), 4.10 (2H, q, *J* 7.1, OC*H*$_2$), 7.12-7.32 (10H, m, 10 × Ar-H); δ_C (50 MHz; $CDCl_3$) 14.4 (*C*H$_3$), 52.8 (N*C*H$_2$), 59.1 (N*C*H$_2$Ar), 60.9 (O*C*H$_2$CH$_3$), 68.9 (O*C*H$_2$COO), 70.1 (O*C*H$_2$), 70.5 (O*C*H$_2$), 71.1 (O*C*H$_2$), 126.9 (Ar-*C*H), 128.3 (Ar-*C*H), 128.9 (Ar-*C*H), 139.8 (Ar-*C*), 170.6 (*C*OO); *m/z* (ESI) 372 ([M+H]$^+$, 100%), 153 (50); HRMS (ESI) (Found: [M+H]$^+$, 372.2170. $C_{22}H_{29}NO_4$ requires *m/z*, 372.2169).

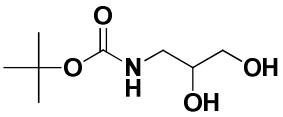
Preparation of *N,N*-dibenzyl-8-amino-3,6-dioxaoctanoic acid (45)[126]



$C_{20}H_{25}NO_4$

Ester **59** (1.57 g, 4.23 mmol) was dissolved in methanol (15 mL) and 1 M sodium hydroxide solution (14 mL) was added to this clear solution. The mixture was stirred at room temperature for 3 h. After this time, the pH of the solution was adjusted to 7 with the addition of 6 M hydrochloric acid solution and methanol was evaporated under reduced pressure. The aqueous solution was further acidified to pH 6 by the addition of 6 M hydrochloric acid solution and extracted with ethyl acetate (6 × 20 mL). The organic phases were separated and dried over anhydrous magnesium sulphate. Evaporation of the solvents yielded a crude residue that required flash column chromatography (ethyl acetate-methanol-glacial acetic acid 89:10:1). The title compound (**45**; 0.800 g, 2.33 mmol; 55%) was obtained as a colourless oil; R_f (ethyl acetate-methanol-glacial acetic acid 89:10:1) = 0.18; υ_{max}(film)/cm^{-1} 3448br w, 3061w, 3028w, 2876m, 1729m, 1599m; δ_H (200 MHz; $CDCl_3$) 2.71 (2H, t, *J* 6.2, NC*H*$_2$), 3.38-3.57 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.77 (4H, s, 2 × NC*H*$_2$Ar), 3.91 (2H, s, OC*H*$_2$CO), 7.12-7.37 (10H, m, 10 × Ar-H), 8.60 (1H, br s, COO*H*); δ_C (50 MHz; $CDCl_3$) 51.9 (N*C*H$_2$), 58.4 (N*C*H$_2$Ar), 68.6 (O*C*H$_2$COO), 70.0 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.5 (O*C*H$_2$), 127.7 (Ar-*C*H), 128.5 (Ar-*C*H), 129.7 (Ar-*C*H), 136.8 (Ar-*C*), 174.2 (*C*OOH).

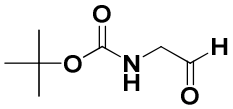
Preparation of *tert*-butoxycarbonylamino-1,2-propanediol (65)[139]



$C_8H_{17}NO_4$

A solution of 3-amino-1,2-propanediol (**64**; 4.96 g, 54.4 mmol) in water (130 mL) was cooled in an ice bath. A solution of di-*tert*-butyl dicarbonate (17.9 g, 82.0 mmol) in dioxane (5 mL) was added and the temperature was allowed to rise to room temperature, whilst the pH was maintained at 10.5 by the addition of 2 M sodium hydroxide solution. After 2 h, this mixture was concentrated *in vacuo* and the paste was triturated with dichloromethane (175 mL). The suspension was then filtered and the organic phase was dried over anhydrous magnesium sulphate. After filtration, the solvent was removed to afford the title compound (**65**; 9.31 g, 48.7 mmol; 90%) as an oil which solidified on cooling to a white solid; R_f (dichloromethane-ethyl acetate 1:1) = 0.60; υ_{max}(film)/cm^{-1} 3358br s, 2979s, 2934s, 2520s, 1792w, 1745m, 1690s, 1526s; δ_H (200 MHz; $CDCl_3$) 1.40 (9H, s, C(C*H*$_3$)$_3$), 2.95-3.05 (2H, m, NHC*H*$_2$CH), 3.22-3.45 (2H, m, C*H*$_2$OH), 3.49-3.53 (1H, m, C*H*), 3.82 (2H, br s, 2 × O*H*), 5.07 (1H, br t, N*H*); δ_C (50 MHz; $CDCl_3$) 28.3 (C(*C*H$_3$)$_3$), 42.7 (NH*C*H$_2$), 63.5 (*C*H$_2$OH), 71.3 (*C*H), 80.2 (*C*(CH$_3$)$_3$), 157.4 (*C*ONH).

Preparation of (*tert*-butoxycarbonyl)aminoacetaldehyde (66)[139]

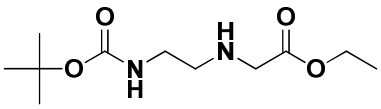


$C_7H_{13}NO_3$

Sodium periodate (9.31 g, 48.7 mmol) was added to *tert*-butoxycarbonylamino-1,2-propanediol (**65**; 10.7 g, 50.2 mmol) in water (100 mL) and the solution was stirred at room temperature for 2 h. The mixture was then filtered and the filtrate was extracted with dichloromethane (5 × 100 mL). The combined organic phase was dried over anhydrous magnesium sulphate and after filtration, the solvent was removed to yield the title compound (**66**; 7.13 g, 44.8 mmol; 92%) as a colourless oil that solidified on cooling; R_f (dichloromethane-methanol 9:1) = 0.61; υ_{max}(film)/cm^{-1} 2979s, 2935s, 1706br s, 1522br

s, 1455s; δ_H (200 MHz; $CDCl_3$) 1.43 (9H, s, C(CH_3)$_3$), 4.06 (2H, d, *J* 5.0, NHCH_2), 5.23 (1H, br s, N*H*), 9.63 (1H, apparent s, C*H*O); δ_C (50 MHz; $CDCl_3$) 28.2 (C(CH_3)$_3$), 51.3 (CH_2), 80.1 ($C(CH_3)_3$), 155.2 (*C*ONH), 197.2 (*C*HO).

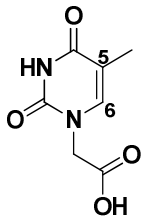
Preparation of ethyl *N*-[2-(*tert*-butoxycarbonyl)aminoethyl]glycinate (67)[139]



$C_{11}H_{22}N_2O_4$

To a solution of (*tert*-butoxycarbonylamino)acetaldehyde (**66**; 7.13 g, 44.8 mmol) in methanol (100 mL) was added glycine ethyl ester hydrochloride (15.6 g, 112 mmol) and sodium borohydride (2.81 g, 44.8 mmol). The mixture was stirred at room temperature for 17 h. After this time, the solvent was evaporated and the residue was dissolved in water (120 mL). The pH of the resulting aqueous solution was adjusted to 8 by addition of 2 M sodium hydroxide solution. This solution was then extracted with dichloromethane (5 × 170 mL). The combined organic fraction was washed with brine (50 mL), dried over anhydrous magnesium sulphate and, after filtration, the solvent was evaporated to give a crude colourless oil. This oil was subjected to flash column chromatography (ethyl acetate–methanol 95:5) to afford the title compound (**67**; 6.03 g, 24.5 mmol; 55%) as a colourless oil; R_f (ethyl acetate-methanol 95:5) = 0.41; υ_{max}(film)/cm^{-1} 2978m, 2933m, 1739s, 1713s, 1518m, 1457w; δ_H (200 MHz; $CDCl_3$) 1.15 (3H, t, *J* 7.1, CH_3), 1.31 (9H, s, (CH_3)$_3$), 1.77 (1H, br s, CH_2N*H*CH_2), 2.61 (2H, t, *J* 5.8, NHCH_2), 3.08 (2H, br dt, CH_2CH_2NH), 3.27 (2H, s, CH_2CO), 4.05 (2H, q, *J* 7.1, OCH_2), 5.32 (1H, br t, N*H*CH_2); δ_C (50 MHz; $CDCl_3$) 13.9 (*C*H_3), 28.1 ((*C*H_3)$_3$), 39.9 (NH*C*H_2), 48.5 (CH_2*C*H_2NH), 50.1 (*C*H_2CO), 60.5 (O*C*H_2), 78.7 (*C*(CH_3)$_3$), 155.9 (*C*ONH), 172.2 (*C*OO).

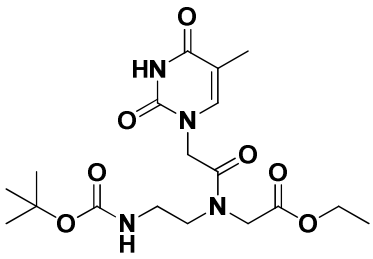
Preparation of thymin-1-yl acetic acid (69)[140]



$C_7H_8N_2O_4$

Thymine (**68**; 3.79 g, 31.6 mmol) was added to a solution of potassium hydroxide (6.74 g, 120 mmol) in water (20 mL). The solution was warmed in a water bath to 40 °C whereupon a solution of bromoacetic acid (6.59 g, 47.6 mmol) in water (10 mL) was slowly added over 30 min. The resulting mixture was stirred for another 30 min before being cooled to room temperature. The pH was then adjusted to 5.5 using concentrated hydrochloric acid. The reaction mixture was then cooled at 4 °C for 2 h. The following solution was then filtered and the pH was further reduced to 2.0 using concentrated hydrochloric acid before being cooled at -20 °C for 2 h. The resulting precipitate was collected by filtration and dried *in vacuo* over P_2O_5 to obtain the title compound as a white solid (**69**; 5.64 g, 30.6 mmol; 97%); R_f (ethyl acetate-methanol-acetic acid 75:20:5) = 0.21; υ_{max}(KBr)/cm^{-1} 3410s, 3181s, 2962s, 2520s, 1739s, 1707s, 1665s, 1634s, 1485s; δ_H (200 MHz; DMSO-d_6) 1.75 (3H, s, C*H*$_3$), 4.40 (2H, s, C*H*$_2$), 7.60 (1H, s, C(6)H); δ_C (50 MHz; DMSO-d_6) 1.75 (*C*H$_3$), 48.9 (*C*H$_2$), 108.8 (*C*(5)), 142.3 (*C*(6)H), 151.4 (*C*(2)O), 164.9 (*C*(4)O), 170.1 (*C*OOH).

Preparation of N-(2-tert-butoxycarbonylaminoethyl)-N-(thymin-1-ylacetyl) glycine ethyl ester (70)[139]

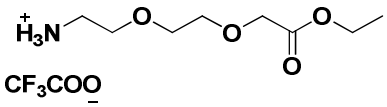


$C_{18}H_{28}N_4O_7$

To thymin-1-ylacetic (**69**; 0.392 g, 2.13 mmol) in dimethylformamide (5 mL) was added HBTU (0.803 g, 2.12 mmol) followed by triethylamine (0.30 mL, 2.2 mmol) and the mixture was stirred for 15 min. After this time, a solution of ethyl N-[2-(tert-butoxycarbonyl)aminoethyl]glycinate (**67**; 0.478 g, 1.94 mmol) in dimethylformamide (5 mL) was added to this mixture followed by triethylamine (0.57 mL, 4.1 mmol). The resulting mixture was stirred for 17 h at room temperature after which time the solvent was removed in vacuo. The residue was dissolved in dichloromethane (100 mL) and the solution was washed successively with 1 M sodium bicarbonate (3 × 50 mL), 1 M potassium hydrogensulphate (2 × 50 mL), water (50 mL) and finally brine (50 mL). The organic layer was then dried over anhydrous magnesium sulphate and the solvent was evaporated under reduced pressure to yield a cream-coloured foam which was purified by flash column chromatography (dichloromethane-methanol 95:5) yielding the title compound (**70**; 0.615 g, 1.49 mmol; 70%) as a white foam; R_f (dichloromethane-methanol 95:5) = 0.44; υ_{max}(KBr)/cm^{-1} 3347m, 3193m, 2980m, 2972m, 1678br s, 1531m; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers in a 2:1 ratio observed due to restricted rotation about the tertiary amide bond) 1.29 (ma.) and 1.34 (mi.) (3H, t, *J* 7.1, C*H*$_3$), 1.44 (ma.) and 1.46 (mi.) (9H, s, (C*H*$_3$)$_3$), 1.93 (ma.) and 1.94 (mi.) (3H, s, C(5)-C*H*$_3$), 3.27-3.30 (mi.) and 3.34-3.36 (ma.) (2H, m, NHC*H*$_2$CH$_2$), 3.49-3.55 (2H, m, NHCH$_2$C*H*$_2$), 4.07 (ma.) and 4.22 (mi.) (2H, s, C*H*$_2$COO), 4.22 (ma.) and 4.28 (mi.) (2H, q, *J* 7.1, OC*H*$_2$), 4.45 (mi.) and 4.60 (ma.) (2H, s, C*H*$_2$CON), 5.10 (mi.) and 5.74 (ma.) (1H, br t, *J* 5.8, N*H*), 7.00 (mi.) and 7.05 (ma.) (1H, s, C(6)H), 9.30 (ma.) and 9.37 (mi.) (1H, br s, N(3)H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers in a 2:1 ratio observed due to restricted rotation about the tertiary amide bond) 12.3 (C(5)-*C*H$_3$), 14.0 (*C*H$_3$), 28.3 (mi.) and 28.4 (ma.) ((*C*H$_3$)$_3$), 38.6 (ma.) and 38.7 (mi.) (NH*C*H$_2$CH$_2$), 47.6 (mi.) and 47.7 (ma.) (*C*H$_2$CON), 48.8 (NHCH$_2$*C*H$_2$), 49.1 (ma.) and 50.2 (mi.)

(*C*H_2COO), 61.7 (ma.) and 62.2 (mi.) (O*C*H_2), 79.9 (*C*(CH_3)$_3$), 110.7 (ma.) and 110.8 (mi.) (C(5)), 140.9 (C(6)H), 151.1 (ma.) and 151.2 (mi.) (C(2)O), 156.0 (*C*ONH), 164.2 (C(4)O), 167.3 (ma.) and 167.7 (mi.) (*C*ON), 169.3 (mi.) and 169.6 (ma.) (*C*OO). The identity of the product was verified by COSY and ^{15}N-^{1}H multiple-bond correlation. An X-Ray crystal structure (Appendix A) was acquired from crystals obtained from diethyl ether and methanol, thereby confirming the structure of the product.

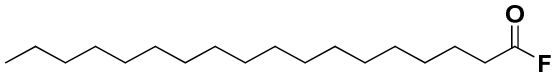
Preparation of 8-amino-3,6-dioxaoctanoic acid ethyl ester trifluoroacetate (77)



$C_{10}H_{18}F_3NO_6$

To a solution of 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctanoic acid ethyl ester (**47**; 0.10 g, 0.34 mmol) in dichloromethane (2 mL) was slowly added trifluoroacetic acid (2 mL). The reaction mixture was stirred for 1 h at room temperature after which time the solvents were removed. The brown residue obtained was coevaporated with toluene (3 × 5 mL) and diethyl ether (3 × 5 mL) and used without further purification.

Preparation of stearoyl fluoride (78)[2]

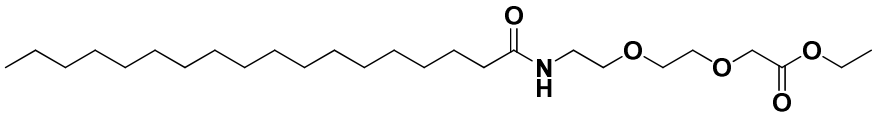


$C_{18}H_{35}FO$

To a solution of stearic acid (**75**; 0.711 g, 2.50 mmol) in dichloromethane (10 mL) was added pyridine (0.202 mL, 2.50 mmol) followed by cyanuric fluoride (0.432 mL, 5.00 mmol). The mixture was heated to reflux for 3 h and after which time dichloromethane (10 mL) was added. The solution was filtered and washed with water (2 × 20 mL), the organic phase was then dried over anhydrous magnesium sulphate. The solvent was finally removed under reduced pressure to yield the title compound (**78**; 0.658 g, 2.30 mmol; 92%) as a white oily solid; υ_{max}(film)/cm^{-1} 2924s, 2854s, 1846s; δ_H (200 MHz; $CDCl_3$) 0.88 (3H, t, *J* 6.3, C*H*$_3$), 1.26 (28H, s, 14 × C*H*$_2$), 1.67 (2H, m, C*H*$_2$CH$_2$COF), 2.49 (2H, dt, $^3J_{HH}$ 7.5, $^3J_{HF}$ 1.3, C*H*$_2$COF); δ_C (50 MHz; $CDCl_3$) 14.0

(*C*H_3), 22.6 (*C*H_2), 23.9 (*C*H_2), 28.6 (*C*H_2), 28.9 (*C*H_2), 29.3 (*C*H_2), 29.6 (*C*H_2), 31.6 (*C*H_2), 31.9 (*C*H_2), 32.6 (*C*H_2COF) and 163.8 (d, *C*OF, $^1J_{13C\text{-}19F}$ 361). Seven signals corresponding to *C*H_2 are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

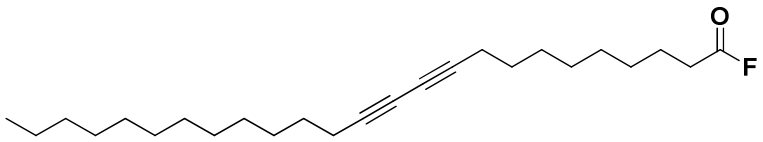
Preparation of ethyl 8-stearamido-3,6-dioxaoctanoate (79)



$C_{26}H_{51}NO_5$

To a solution of **77** (0.10 g, 0.33 mmol) in dimethylformamide (5 mL) was added DIPEA (0.287 mL; 1.65 mmol) followed by a solution of stearoyl fluoride (**78**; 0.13 g, 0.46 mmol) in dichloromethane (5 mL). The mixture was stirred for 20 h at room temperature. After this time the solvent was removed under reduced pressure and the residue was dissolved in dichloromethane (10 mL). This was then washed with 1 M sodium bicarbonate (2 × 10 mL), 1 M potassium hydrogen sulphate (2 × 10 mL), water (15 mL) and brine (15 mL). The organic layer was dried over magnesium sulphate and the solvent removed under reduced pressure. The resulting yellow solid was subjected to flash column chromatography (ethyl acetate-methanol 95:5) to yield the title compound (**79**; 0.12 g, 0.26 mmol, 79%) as a white solid: m.p. 65-66 °C; R_f (ethyl acetate-methanol 95:5) = 0.25; (Found C, 68.14; H, 11.17; N, 3.21% $C_{26}H_{51}NO_5$ requires C, 68.23; H, 11.23; N, 3.06%); υ_{max}(film)/cm^{-1} 3307s, 3083m, 2980s, 2850s, 1749s, 1703s, 1641s; δ_H (200 MHz, $CDCl_3$) 0.68 (3H, t, *J* 6.3, C*H*$_3$), 0.95-1.26 (31H, m, OCH$_2$C*H*$_3$ and 14 × C*H*$_2$), 1.31-1.48 (2H, m, C*H*$_2$CH$_2$CO), 1.97 (2H, br t, *J* 7.5, C*H*$_2$CONH), 3.22 (2H, br dt, NHC*H*$_2$), 3.30-3.53 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.40 (2H, s, OC*H*$_2$CO), 4.01 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 5.95 (1H, br t, N*H*); δ_C (50 MHz, $CDCl_3$) 13.6 (*C*H$_3$CH$_2$ and OCH$_2$*C*H$_3$), 22.1 (*C*H$_2$), 25.2 (*C*H$_2$CH$_2$CO), 28.8 (*C*H$_2$), 28.9 (*C*H$_2$), 29.0 (*C*H$_2$), 29.1 (*C*H$_2$), 29.2 (*C*H$_2$), 31.4 (*C*H$_2$), 36.2 (*C*H$_2$CO), 38.5 (NH*C*H$_2$), 60.4 (O*C*H$_2$CH$_3$), 68.0 (O*C*H$_2$CO), 69.5 (O*C*H$_2$), 69.6 (O*C*H$_2$), 70.4 (O*C*H$_2$), 160.3 (*C*OO), 173.2 (*C*ONH); *m/z* (EI) 458 ([M]$^+$, 22%), 327 (80), 234 (74), 132 (96), 103 (90), 85 (100). Seven signals corresponding to *C*H_2 from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

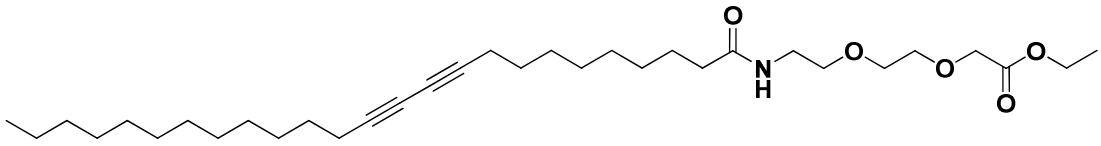
Preparation of pentacosa-10,12-diynoyl fluoride (80)[2]



$C_{25}H_{41}FO$

To a solution of 10,12-pentacosadiynoic acid (**5**; 0.716 g, 1.91 mmol) in dichloromethane (20 mL) was added pyridine (0.155 mL, 1.92 mmol) followed by cyanuric fluoride (0.230 mL, 2.67 mmol). This mixture was heated to reflux for 3 h after which time dichloromethane (10 mL) was added and the mixture filtered. The filtrate was then washed with water (2 × 20 mL) and the organic extract was dried over anhydrous magnesium sulphate. The solvent was evaporated under reduced pressure to yield the title compound (**80**; 0.589 g, 1.56 mmol; 82 %) as a white oily solid; υ_{max}(film)/cm^{-1} 2917m, 2845m, 2240w, 1863s; δ_H (200 MHz; $CDCl_3$) 0.85 (3H, t, *J* 6.4, C*H*$_3$), 1.20-1.40 (30H, m, 15 × C*H*$_2$), 1.51-1.72 (2H, m, C*H*$_2$CH$_2$COF), 2.15 (4H, t, *J* 6.7, 2 × C*H*$_2$C≡C), 2.45 (2H, dt, $^3J_{1H\text{-}1H}$ 4.9, $^3J_{1H\text{-}19F}$ 0.9, C*H*$_2$COF); δ_C (50 MHz; $CDCl_3$) 14.2 (*C*H$_3$), 19.2 (*C*H$_2$C≡C), 22.7 (*C*H$_2$), 23.9 (*C*H$_2$), 28.3 (*C*H$_2$), 28.9 (*C*H$_2$), 28.7 (*C*H$_2$), 28.8 (*C*H$_2$), 28.9 (*C*H$_2$), 28.9 (*C*H$_2$), 29.2 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.7 (*C*H$_2$), 31.6 (*C*H$_2$), 31.9 (*C*H$_2$), 32.6 (*C*H$_2$COF), 65.2 and 65.4 (C≡*C*-*C*≡C), 77.3 and 77.7 (*C*≡C-C≡*C*), 160.3 (d, *C*OF, $^1J_{13C\text{-}19F}$ 361). Three signals corresponding to *C*H$_2$ are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of ethyl 8-(pentacosa-10,12-diynamido)-3,6-dioxactonoate (81)

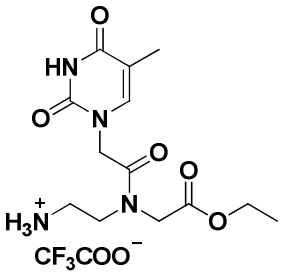


$C_{33}H_{57}NO_5$

To a solution of **77** (0.058 g, 0.19 mmol) in dimethylformamide (5 mL) was added DIPEA (0.17 mL, 0.95 mmol) followed by a solution of pentacosa-10,12-diynoyl fluoride (**80**; 0.087 g, 0.23 mmol) in dichloromethane (5 mL). The mixture was stirred for 20 h at room temperature. After this time the solvent was removed under reduced pressure and the residue was dissolved in dichloromethane (10 mL). This was then washed

with 1 M sodium bicarbonate (2 × 10 mL), 1 M potassium hydrogen sulphate (2 × 10 mL), water (15 mL) and brine (15 mL). The organic layer was dried over magnesium sulphate and the solvent removed under reduced pressure. The resulting yellow solid was subjected to flash column chromatography (dichloromethane-methanol 95:5) to yield the title compound (**81**; 0.072 g, 0.13 mmol, 69 %) as a white solid: m.p. 60 °C; δ_H (200 MHz, $CDCl_3$) 0.93 (3H, t, *J* 6.4, C*H*$_3$CH$_2$), 1.23-1.76 (35H, m, OCH$_2$C*H*$_3$ and 16 × C*H*$_2$), 2.15 (2H, t, *J* 7.6, C*H*$_2$CONH), 2.28 (4H, t, *J* 6.7, 2 × C*H*$_2$C≡C), 3.52 (2H, br dt, NHC*H*$_2$CH$_2$O), 3.58-3.66 (2H, m, NHCH$_2$C*H*$_2$O), 3.68-3.72 (m, 4H, OC*H*$_2$C*H*$_2$O), 4.19 (2H, s, OC*H*$_2$CO), 4.28 (2H, q, *J* 7.1, OC*H*$_2$), 6.29 (1H, br t, N*H*); δ_C (50 MHz, $CDCl_3$) 13.8 (C*H*$_3$), 14.1 (OCH$_2$C*H*$_3$), 18.9 (C*H*$_2$C≡C), 22.4 (C*H*$_2$), 25.4 (C*H*$_2$), 28.1 (C*H*$_2$), 28.5 (C*H*$_2$), 28.6 (C*H*$_2$), 28.7 (C*H*$_2$), 28.8 (C*H*$_2$), 28.9 (C*H*$_2$), 29.0 (C*H*$_2$), 29.2 (C*H*$_2$), 29.3 (C*H*$_2$), 31.6 (C*H*$_2$), 36.4 (C*H*$_2$CO), 38.8 (NHC*H*$_2$), 60.7 (OC*H*$_2$CH$_3$), 64.9 and 65.2 (C≡*C*-*C*≡C), 68.2 (O*C*H$_2$CO), 69.7 (O*C*H$_2$), 69.8 (O*C*H$_2$), 70.7 (O*C*H$_2$), 77.2 and 77.3 (*C*≡C-C≡*C*), 170.2 (*C*OO), 173.0 (*C*ONH); *m/z* (ESI) 570 ([M+Na]$^+$, 45%), 378 (13), 101 (40), 81 (33), 59 (100); HRMS (ESI) (Found: [M+Na]$^+$, 570.4125. $C_{33}H_{57}N\,O_5Na$ requires *m/z*, 570.4129). Five signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of *N*-(2-aminoethyl)-*N*-(thymin-1-ylacetyl)glycine ethyl ester trifluoroacetate (82)

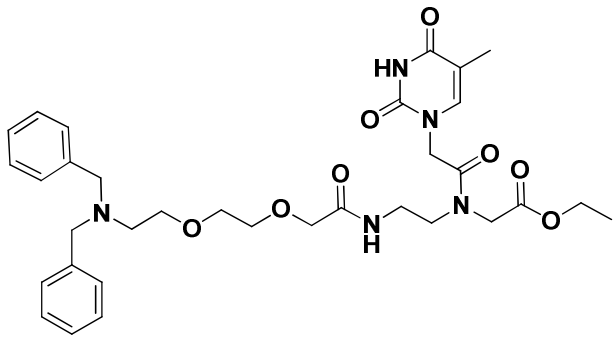


$C_{15}H_{21}F_3N_4O_7$

PNA monomer **70** (0.0536 g, 0.130 mmol) was dissolved in dichloromethane (2 mL) and trifluoroacetic acid (2 mL) was added dropwise. The solution was stirred for 1 h after which time the solvent was evaporated under reduced pressure to yield a brown residue. This residue was coevaporated with toluene (2 × 5 mL) and diethyl ether (4 × 5 mL) and finally dried *in vacuo* over phosphorus pentoxide, to yield a brown sticky solid (**82**; 0.0554 g, 0.130 mmol; 100%) which was used without further purification; δ_H

(200 MHz; D_2O) (2 rotational isomers in a 2:1 ratio observed due to restricted rotation about the tertiary amide bond) 1.10 (mi.) and 1.15 (ma.) (3H, t, *J* 7.1, C*H*$_3$), 1.74 (3H, s, C(5)-C*H*$_3$), 3.02-3.15 (ma.) and 3.18-3.29 (mi.) (2H, m, NHC*H*$_2$CH$_2$), 3.58-3.68 (ma.) and 3.70-3.80 (mi.) (2H, m, NHCH$_2$C*H*$_2$), 4.06 (mi.) and 4.28 (ma.) (2H, s, C*H*$_2$COO), 4.09 (mi.) and 4.17 (ma.) (2H, q, *J* 7.1, OC*H*$_2$), 4.51 (ma.) and 4.69 (mi.) (2H, s, C*H*$_2$CON), 7.20 (ma.) and 7.26 (mi.) (1H, s, C(6)H); δ_C (50 MHz; D_2O) (2 rotational isomers in a 2:1 ratio observed due to restricted rotation about the tertiary amide bond) 10.7 (C(5)-*C*H$_3$), 12.7 (*C*H$_3$), 37.1 (NH*C*H$_2$CH$_2$), 48.5 (*C*H$_2$CON), 48.8 (NHCH$_2$*C*H$_2$), 49.0 (*C*H$_2$COO), 62.7 (O*C*H$_2$), 110.6 (C(5)), 142.3 (C(6)H), 151.7 (C(2)O), 166.1 (C(4)O), 169.1 (*C*ON), 170.2 (*C*OO). Signals corresponding to NH were exchanged on deuteration.

Attempted preparation of ethyl 14-(dibenzylamino)-3-(2-(thymin-1-yl)acetyl)-7-oxo-9,12-dioxa-3,6-diazatetradecan-1-oate (76)

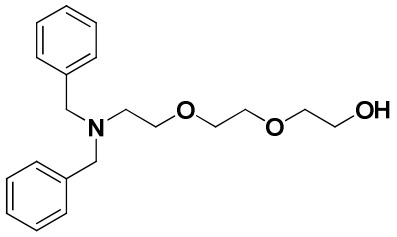


$C_{33}H_{43}N_5O_8$

To a solution of **45** (0.33 g, 0.80 mmol) in dimethylformamide (5 mL) was added HBTU (0.31 g, 0.81 mmol) followed by diisopropylethylamine (0.15 mL, 0.83 mmol) and the mixture was stirred for 15 min. After this time, a solution of **82** (0.33 g, 0.78 mmol) in dimethylformamide (5 mL) was added to this mixture followed by diisopropylethylamine (0.30 mL, 1.66 mmol). The resulting mixture was stirred for 18 h at room temperature after which time the solvent was removed *in vacuo*. The residue was dissolved in dichloromethane (10 mL) and the solution was washed successively with 1 M sodium bicarbonate (3 × 10 mL), 1 M potassium hydrogensulphate (2 × 10 mL), water (10 mL) and finally brine (15 mL). The organic layer was then dried over anhydrous magnesium sulphate and the solvent was evaporated under reduced pressure to yield a yellow residue which was purified by flash column chromatography (dichloromethane-

methanol 95:5) yielding the title compound (**76**; 0.029 g, 0.045 mmol; 6%) as a yellow sticky powder. The yield was too low for recording a clean ^{1}H NMR.

Preparation of *N*,*N*-dibenzyl-8-amino-3,6-dioxaoctan-1-ol (100)



$C_{20}H_{27}NO_3$

Method A

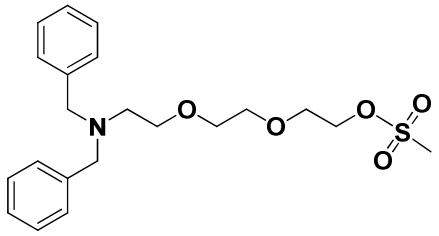
Lithium aluminium hydride (0.0590 g, 1.58 mmol) was dissolved in anhydrous tetrahydrofuran (1.60 mL) and the slurry obtained was cooled to -78 °C. Compound **59** (0.590 g, 1.58 mmol) in anhydrous tetrahydrofuran (20 mL) was added dropwise. The mixture was stirred for 1 h, after being allowed to slowly warm-up to room temperature. It was left stirring for another 17 h. After this time 2 M ammonium hydroxide solution (1 mL) was added carefully and the solvent was evaporated. The aqueous phase was extracted with ethyl acetate (5 × 20 mL) and the combined organic layer was dried over anhydrous magnesium sulphate. The filtrate was concentrated and the crude material was subjected to flash column chromatography (ethyl acetate - light petroleum 3:2) to yield the title compound (**100**; 0.406 g, 1.23 mmol; 72%) as a colourless oil.

Method B

Compound **59** (5.83 g, 15.7 mmol) in tetrahydrofuran (80 mL) was heated to reflux and to this was added sodium borohydride (3.86 g, 102 mmol) over a period of 15 min. Methanol (80 mL) was then carefully added dropwise to the refluxing reaction over another 15 min. The reaction mixture finally obtained was further heated to reflux for 60 min. After this time, the reaction was quenched by the addition of saturated ammonium chloride solution (80 mL) and the organic phase was separated. The aqueous phase was extracted with ethyl acetate (6 × 80 mL) and the combined organic phase was dried over anhydrous magnesium sulphate. The filtrate was concentrated to a crude oil, which was then subjected to flash column chromatography (ethyl acetate – light petroleum 2:1) to yield the title compound (**100**; 4.22 g, 12.8 mmol; 82%) as a colourless oil, a minor

product (**43**; 0.896 g, 3.14 mmol; 20%) was also isolated as a colourless oil; R_f (ethyl acetate-light petroleum 7:3) = 0.33; υ_{max}(film)/cm^{-1} 3436br s, 3027s, 2871s, 2871m, 1494s, 1453s; δ_H (200 MHz; $CDCl_3$) 2.73 (2H, t, *J* 6.2, NC*H*$_2$), 2.90 (1H, br s, O*H*), 3.57-3.69 (10H, m, C*H*$_2$OC*H*$_2$C*H*$_2$OC*H*$_2$C*H*$_2$), 3.72 (4H, s, 2 × NC*H*$_2$Ar), 7.24-7.48 (10H, m, 10 × Ar-H); δ_C (50 MHz; $CDCl_3$) 52.8 (N*C*H$_2$), 59.1 (2 × N*C*H$_2$Ar), 61.8 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.5 (2 × O*C*H$_2$), 72.7 (O*C*H$_2$), 127.0 (Ar-*C*H), 128.3 (Ar-*C*H), 128.9 (Ar-*C*H), 139.8 (Ar-*C*); *m/z* (CI) 330 ([M+H]$^+$, 100%), 268 (3), 254 (6), 240 (30), 238 (23), 210 (30); HRMS (ESI) (Found: [M+H]$^+$, 330.2068. $C_{20}H_{27}NO_3$ requires *m/z*, 330.2064).

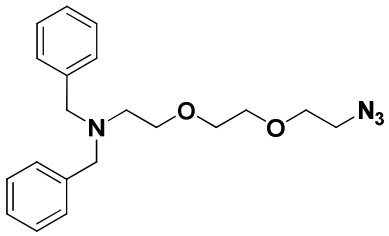
Preparation of *N,N*-dibenzyl-8-amino-3,6-dioxaoctanyl methanesulfonate (101)



$C_{21}H_{29}NO_5S$

To an ice-cooled solution of **100** (0.665 g, 2.02 mmol) and triethylamine (0.430 mL, 3.09 mmol) in anhydrous dichloromethane (10 mL) was slowly added *via* a syringe methanesulfonyl chloride (0.170 mL, 2.22 mmol). The reaction was stirred and allowed to slowly warm up to room temperature. The stirring was continued overnight and after this time the reaction mixture was washed successively with ice cooled water (15 mL), saturated sodium bicarbonate solution (15 mL) and brine (20 mL). The organic phase was dried over magnesium sulphate and evaporated. The crude residue obtained was subjected to flash column chromatography (ethyl acetate - light petroleum 4:6) and the title compound (**101**; 0.675 g, 1.66 mmol; 82%) was obtained as a colourless liquid; R_f (ethyl acetate-light petroleum 4:6) = 0.29; υ_{max}(film)/cm^{-1} 3027s, 2871s, 2871m, 1494s, 1453s, 1354s (SO_2OR str as), 1176s (SO_2OR str sy); δ_H (200 MHz; $CDCl_3$) 2.62 (2H, t, *J* 6.2, NC*H*$_2$), 2.92 (3H, s, SO_2C*H*$_3$), 3.45-3.56 (6H, m, OC*H*$_2$C*H*$_2$OC*H*$_2$), 3.59 (4H, s, 2 × NC*H*$_2$Ar), 3.63-3.68 (2H, m, NCH$_2$C*H*$_2$O), 4.25-4.29 (2H, m, C*H*$_2$OSO$_2$), 7.13-7.35 (10H, m, 10 × Ar-H); δ_C (50 MHz; $CDCl_3$) 37.6 (SO_2*C*H$_3$), 52.6 (N*C*H$_2$), 58.9 (2 × N*C*H$_2$Ar), 69.0 (*C*H$_2$SO$_2$), 69.3 (*C*H$_2$CH$_2$SO$_2$), 70.0 (O*C*H$_2$), 70.2 (O*C*H$_2$), 70.6 (O*C*H$_2$), 126.9 (Ar-*C*H), 128.2 (Ar-*C*H), 128.7 (Ar-*C*H), 139.6 (Ar-*C*).

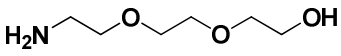
Preparation of *N*,*N*-dibenzyl-8-amino-3,6-dioxaoctan-1-azide (102)



$C_{20}H_{26}N_4O_2$

Compound **101** (1.05 g, 2.59 mmol) was dissolved in anhydrous dimethylformamide (25 mL) and sodium azide (0.42 g, 6.5 mmol) was then added in one portion. The reaction mixture was gently heated to 60 °C for 4 h. After this time the mixture was poured on a mixture of ice/water (*ca.* 20 g) and extracted with dichloromethane (3 × 50 mL). The organic phase was separated, dried over magnesium sulphate and the solvent was removed to yield the title compound (**102**; 0.798 g, 2.25 mmol; 87%) as a yellow liquid; υ_{max}(film)/cm^{-1} 3061m, 3027s, 2977s, 2827s, 2799s, 2104s (N_3), 1680m; δ_H (200 MHz; $CDCl_3$) 2.77 (2H, t, *J* 6.2, NC*H*$_2$), 3.41 (3H, t, *J* 5.1, C*H*$_2$N$_3$), 3.60-3.76 (8H, m, C*H*$_2$OC*H*$_2$C*H*$_2$OC*H*$_2$), 3.71 (4H, s, 2 × NC*H*$_2$Ar), 7.21-7.49 (10H, m, 10 × Ar-H); δ_C (50 MHz; $CDCl_3$) 50.5 (*C*H$_2$N$_3$), 52.5 (N*C*H$_2$), 58.8 (2 × N*C*H$_2$Ar), 69.8 (O*C*H$_2$), 69.9 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.5 (O*C*H$_2$), 126.7 (Ar-*C*H), 127.9 (Ar-*C*H), 128.6 (Ar-*C*H), 139.6 (Ar-*C*); *m/z* (CI) 355 ([M+H]$^+$, 100%); HRMS (ESI) (Found: [M+H]$^+$, 355.2127. $C_{20}H_{26}N_4O_2$ requires *m/z*, 355.2129).

Preparation of 8-amino-3,6-dioxaoctan-1-ol (103)



$C_6H_{15}NO_3$

Method A

Compound **100** (1.30 g, 3.94 mmol) was dissolved in absolute ethanol (50 mL) and 10% Pd/C (0.2 g) was then carefully added. The black heterogeneous mixture was evacuated and purged with H_2(g) several times and stirred at room temperature under 1 atm of H_2(g) for 48 h. The reaction mixture was then purged with nitrogen and filtered over Celite®, washing with ethanol (100 mL) and ethyl acetate (100 mL). The filtrate was

concentrated to afford the title compound (**103**; 0.58 g, 3.9 mmol; 99%) as a colourless oil.

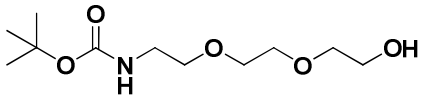
Method B

Compound **100** (1.20 g, 3.64 mmol) was dissolved in methanol (20 mL). Ammonium formate (0.459 g, 7.28 mmol) and 10% Pd/C (1 g) were then carefully added. The black heterogeneous mixture was heated to gentle reflux for 3 h and then filtered over Celite®, washing with methanol (40 mL) and ethyl acetate (40 mL). The filtrate was concentrated to afford the title compound (**103**; 0.533 g, 3.57 mmol; 98%) as a colourless oil.

Method C

Compound **100** (1.00 g, 3.04 mmol) was dissolved in absolute ethanol (60 mL) so that the concentration was 0.05 M. The solution was pumped (1 mL/min) through an H-cube® instrument (provided by Prof Adams) using 10% Pd/C cartridge at 70 bar pressure and at 70 °C. The solution was run two times to ensure completion (using the same cartridge). The solvent was then evaporated to yield the title compound (**103**; 0.45 g, 3.00 mmol; 99%) as colourless oil; υ_{max}(film)/cm^{-1} 3365br s, 3312s, 3302s, 2873s, 2854s, 1599m; δ_H (200 MHz; D_2O) 2.88 (2H, t, *J* 5.3, NC*H*$_2$CH$_2$), 3.54-3.72 (10H, m, C*H*$_2$OC*H*$_2$C*H*$_2$OC*H*$_2$C*H*$_2$); δ_C (50 MHz; D_2O) 39.1 (NH$_2$*C*H$_2$), 60.0 (*C*H$_2$OH), 68.6 (O*C*H$_2$), 69.2 (O*C*H$_2$), 69.3 (O*C*H$_2$) 71.4 (O*C*H$_2$). Signals for protons N*H* and O*H* were not observed due to deuteration.

Preparation of 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctan-1-ol (104)

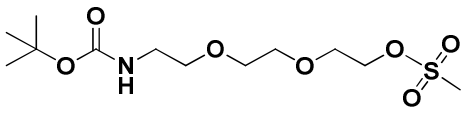


$C_{11}H_{23}NO_5$

Amino alcohol **103** (0.476 g, 3.19 mmol) was dissolved in a solution of 10% triethylamine in methanol (20 mL). Di-*tert*-butyl dicarbonate (1.43 g, 6.57 mmol) was then added in one portion. The reaction mixture was heated to gentle reflux for 3 h, after which time the solvent was removed *in vacuo*. The residual oil was purified by flash column chromatography (ethyl acetate – methanol 95:5) to yield the title compound (**104**; 0.669 g, 2.68 mmol; 84%) as a colourless oil; R_f (ethyl acetate-methanol 95:5) =

0.40; υ_{max}(film)/cm^{-1} 3356br s, 2976s, 1694s, 1526m; δ_H (200 MHz; $CDCl_3$) 1.38 (9H, s, C(C*H*$_3$)$_3$), 2.98 (1H, br s, O*H*), 3.25 (2H, br dt, NHC*H*$_2$), 3.49 (2H, t, *J* 5.3, NHCH$_2$C*H*$_2$O), 3.54-3.61 (6H, m, OC*H*$_2$C*H*$_2$OC*H*$_2$), 3.68 (2H, br dt, CH$_2$C*H*$_2$OH), 5.27 (1H, br s, N*H*); δ_C (50 MHz; $CDCl_3$) 28.7 (C(*C*H$_3$)$_3$), 39.9 (NH*C*H$_2$), 61.3 (*C*H$_2$OH), 69.9 (O*C*H$_2$), 70.5 (O*C*H$_2$), 70.6 (O*C*H$_2$), 72.3 (O*C*H$_2$), 79.5 (*C*(CH$_3$)$_3$), 156.3 (*C*ONH); *m/z* (CI) 250 ([M+H]$^+$, 15%), 150 (100); HRMS (ESI) (Found: [M+H]$^+$, 250.1649. $C_{11}H_{23}NO_5$ requires *m/z*, 250.1649).

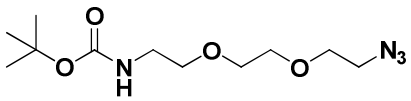
Preparation of 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctanyl methanesulfonate (105)



$C_{12}H_{25}NO_7S$

To an ice-cooled solution of **104** (0.519 g, 2.08 mmol) and triethyethylamine (0.435 mL, 3.12 mmol) in anhydrous dichloromethane (10 mL) was slowly added *via* a syringe methanesulfonyl chloride (0.175 mL, 2.29 mmol). The reaction was stirred and allowed to slowly warm up to room temperature. The stirring was continued overnight and after this time the reaction mixture was washed successively with ice cooled water (10 mL), saturated sodium bicarbonate solution (10 mL) and brine (10 mL). The organic phase was dried over magnesium sulphate and evaporated. The crude residue obtained was subjected to flash column chromatography (ethyl acetate - light petroleum 4:6) and the title compound (**105**; 0.531 g, 1.62 mmol; 78%) was obtained as a colourless liquid; R_f (ethyl acetate-methanol 95:5) = 0.46; υ_{max}(film)/cm^{-1} 2976m, 2935m, 2874m, 1708s, 1521s, 1352s (SO_2OR str as), 1175s (SO_2OR str sy); δ_H (200 MHz; $CDCl_3$) 1.38 (9H, s, C(C*H*$_3$)$_3$), 3.01 (3H, s, SO$_2$C*H*$_3$), 3.24 (2H, br dt, NHC*H*$_2$), 3.47 (2H, t, *J* 5.3, NHCH$_2$C*H*$_2$O), 3.58-3.61 (4H, m, OC*H*$_2$C*H*$_2$O), 3.71 (2H, t, *J* 4.6, OC*H*$_2$CH$_2$SO$_2$), 4.33 (2H, t, *J* 4.6, OCH$_2$C*H*$_2$SO$_2$), 4.93 (1H, br s, N*H*); δ_C (50 MHz; $CDCl_3$) 28.6 (C(*C*H$_3$)$_3$), 37.9 (SO$_2$*C*H$_3$), 40.5 (NH*C*H$_2$), 69.2 (*C*H$_2$SO$_2$), 69.4 (O*C*H$_2$), 70.4 (O*C*H$_2$), 70.5 (O*C*H$_2$), 70.8 (O*C*H$_2$), 79.5 (*C*(CH$_3$)$_3$), 155.8 (*CO*NH); *m/z* (CI) 345 ([M+NH$_3$]$^+$, 17%), 328 ([M+H]$^+$, 20), 289 (70), 271 (100), 228 (35); HRMS (ESI) (Found: [M+H]$^+$, 328.1426. $C_{12}H_{26}NO_7S$ requires *m/z*, 328.1424).

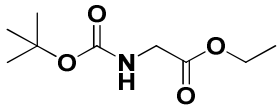
Preparation of 8-(*tert*-butoxycarbonyl)amino-3,6-dioxaoctan-1-azide (83)



$C_{11}H_{22}N_4O_4$

Compound **105** (0.591 g, 1.81 mmol) was dissolved in anhydrous dimethylformamide (20 mL) and sodium azide (0.311 g, 4.76 mmol) was then added in one portion. The reaction mixture was gently heated to 60 °C for 4 h. After this time the mixture was poured on a mixture of ice/water (*ca.* 20 g) and extracted with dichloromethane (3 × 20 mL). The organic phase was separated, dried over magnesium sulphate and the solvent was removed to yield the title compound (**83**; 0.454 g, 1.66 mmol; 92%) as a yellow liquid that did not require any further purification; R_f (ethyl acetate-light petroleum 1:1) = 0.39; υ_{max}(film)/cm^{-1} 3360m, 2977m, 2930m, 2870m, 2109s (N_3), 1713s; δ_H (400 MHz; $CDCl_3$) 1.42 (9H, s, (C*H*$_3$)$_3$), 3.27-3.30 (2H, m, NHC*H*$_2$), 3.38 (2H, t, *J* 5.3, C*H*$_2$N$_3$), 3.52 (2H, t, *J* 5.3, NCH$_2$C*H*$_2$O), 3.59-3.64 (4H, m, OC*H*$_2$C*H*$_2$O), 3.65 (2H, t, *J* 5.3, C*H*$_2$CH$_2$N$_3$), 5.00 (1H, br s, N*H*); δ_C (100 MHz; $CDCl_3$) 28.4 ((*C*H$_3$)$_3$), 40.3 (NH*C*H$_2$), 50.6 (*C*H$_2$N$_3$), 70.0 (O*C*H$_2$CH$_2$N$_3$), 70.2 (O*C*H$_2$CH$_2$O), 70.3 (NCH$_2$*C*H$_2$O), 70.5 (OCH$_2$*C*H$_2$O), 79.1 (*C*(CH$_3$)$_3$), 155.9 (O*C*ONH); *m/z* (CI) 292 ([M+NH$_4$]$^+$, 7%), 275 ([M+H]$^+$, 12), 236 (50), 175 (95), 44 (100); HRMS (ESI) (Found: [M+H]$^+$, 275.1715. $C_{11}H_{23}N_4O_4$ requires *m/z*, 275.1714).

Preparation of *N*-(*tert*-butoxycarbonyl)glycine ethyl ester (110)

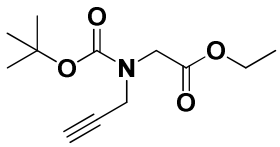


$C_9H_{17}NO_4$

Tert-butyl dicarbonate (5.14 g, 23.5 mmol) in dioxane (7 mL) was added to the stirred solution of glycine ethyl ester hydrochloride (**109**; 3.02 g, 21.6 mmol) dissolved into water (60 mL) and cooled in an ice bath to 0 °C. After 15 min at this temperature, the ice bath was removed and the temperature was allowed to rise to room temperature, whilst the pH was kept just above 10.5 by the addition of 2 M sodium hydroxide solution. Then, the reaction mixture was concentrated to a paste and the paste was triturated

with dichloromethane (50 mL) and water (20 mL). The aqueous phase was separated and extracted one more time with dichloromethane (30 mL). The combined organic phase was dried over magnesium sulphate anhydrous, filtered and the crude material obtained on evaporation was subjected to flash column chromatography (ethyl acetate-light petroleum 6:4). The title compound (**110**; 2.77 g, 13.6 mmol, 63%) was obtained as a yellow liquid; R_f (ethyl acetate-light petroleum 6:4) = 0.66; υ_{max}(film)/cm^{-1} 3372m, 2980m, 2935w, 1732s, 1717s, 1519s; δ_H (200 MHz; $CDCl_3$) 1.22 (2H, t, *J* 7.2, C*H*$_3$), 1.39 (9H, s, (C*H*$_3$)$_3$), 3.83 (2H, d, *J* 5.6, NHC*H*$_2$), 4.1 (2H, q, *J* 7.2, OC*H*$_2$), 5.04 (1H, br s, N*H*CH$_2$); δ_C (50 MHz; $CDCl_3$) 14.0 (*C*H$_3$), 28.2 ((*C*H$_3$)$_3$), 42.3 (NH*C*H$_2$), 61.1 (O*C*H$_2$), 79.6 (*C*(CH$_3$)$_3$), 155.7 (*C*ONH), 170.3 (*C*OO); *m/z* (CI) 204 ([M+H]$^+$, 20%), 165 (90), 104 (100), 58 (10), 44 (16); HRMS (ESI) (Found: [M+H]$^+$, 204.1230. $C_9H_{18}NO_4$ requires 204.1230).

Preparation of *N*-*tert*-butoxycarbonyl-*N*-(prop-2-ynyl)glycine ethyl ester (111)



$C_{12}H_{19}NO_4$

Method A

Ethyl 2-(*tert*-butoxycarbonylamino)acetate (**110**; 0.578 g, 2.84 mmol) was dissolved in anhydrous tetrahydrofuran (5 mL) and 3-bromopropyne (80% wt/v in toluene, 0.950 mL, 8.53 mmol) was added. The solution was cooled to 0 °C in an ice bath and small portions of sodium hydride (60% in mineral oil, 0.175 g, 4.38 mmol) were added. After the evolution of hydrogen had ceased, the deep brown solution was stirred for 18 h at room temperature. After this time, water (5 mL) was added carefully and tetrahydrofuran was removed *in vacuo*. The paste obtained was dissolved in ethyl acetate (10 ml) and the aqueous layer was separated and extracted with ethyl acetate (6 × 10 mL). The combined organic layer was washed with water (2 × 10 mL) and brine (10 mL) and the aqueous layer backextracted with ethyl acetate (10 mL). The combined organic phase was dried over magnesium sulphate anhydrous, filtered and evaporated. The residue obtained was subjected to flash column chromatography (ethyl acetate-light petroleum 1:5) to give the title compound (**111**; 0.345 g, 1.43 mmol, 50%) as deep yellow oil.

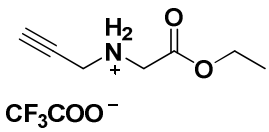
Method B

The same procedure was used as in method A except for the solvent. **110** (0.852 g, 4.19 mmol) was dissolved in anhydrous acetonitrile (4 mL) and 3-bromopropyne (80% wt/v in toluene, 1.40 mL, 12.6 mmol) were then added. Sodium hydride (60% in mineral oil, 0.350 g, 8.75 mmol) was added at 0 °C. The work-up was as described in method A. The title compound (**111**; 0.688 g, 2.85 mmol; 68%) was obtained as deep yellow oil.

Method C

The same procedure was used as in method A except for the solvent. **110** (1.54 g, 7.57 mmol) was dissolved in anhydrous dimethylformamide (5 mL) and 3-bromopropyne (80% wt/v in toluene, 2.53 mL, 22.7 mmol) were then added. Sodium hydride (60% in mineral oil, 0.609 g, 15.2 mmol) was added at 0 °C. The work-up was as described in method A. The title compound (**111**; 1.39 g, 5.76 mmol; 76%) was obtained as deep yellow oil; υ_{max}(film)/cm^{-1} 3270s, 2980s, 2934s, 1748s, 1706s; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.26-1.30 (3H, m, C*H*$_3$), 1.39 and 1.49 (9H, s, (C*H*$_3$)$_3$), 2.22 (1H, t, *J* 2.5, $CH_2C{\equiv}CH$), 3.89 and 4.09 (2H, s, NC*H*$_2$CO), 4.13 and 4.19 (s, 2H, C*H*$_2$CCH), 4.18-4.22 (m, 2H, OC*H*$_2$); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 and 14.2 (*C*H$_3$), 28.1 and 28.2 ((*C*H$_3$)$_3$), 36.6 and 37.1 (*C*H$_2$CCH), 46.9 and 47.5 (*C*H$_2$COO), 61.0 (O*C*H$_2$), 72.3 and 72.6 ($CH_2C{\equiv}CH$), 78.7 (*C*H), 80.9 and 81.0 (*C*(CH$_3$)$_3$), 154.6 and 154.7 (*C*ON), 169.7 (*C*OO); *m/z* (CI) 242 ([M+H]$^+$, 15%), 203 (75), 142 (100), 68 (26); HRMS (ESI) (Found: [M+H]$^+$, 242.1387. $C_{12}H_{19}NO_4$ requires *m/z*, 242.1387). Structure was verified by NOESY experiments.

Preparation of *N*-(prop-2-ynyl)glycine ethyl ester trifluoroacetate (106)

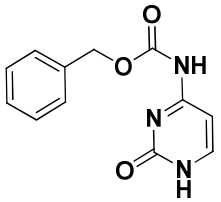


$C_9H_{12}NO_4F_3$

Ethyl 2-*tert*-butoxycarbonyl-*N*-(prop-2-ynyl)aminoacetate (**111**; 1.00 g, 4.14 mmol) was dissolved in dichloromethane (8 mL) and trifluoroacetic acid (8 mL) was added. The

solution was stirred for 1 h. After this time the solvents were evaporated *in vacuo* and the residue was coevaporated with toluene (2 × 10 mL) and diethyl ether (2 × 10 mL) and stored over P_2O_5. The title compound was obtained as a brown oil which was used without further purification (**106**; 1.04 g, 4.06 mmol; 98%); υ_{max}(film)/cm^{-1} 3270w, 2988m, 2930m, 1749s, 1668s; δ_H (200 MHz; $CDCl_3$) 1.29 (3H, t, *J* 7.2, C*H*$_3$), 2.58 (1H, t, *J* 2.3, $CH_2C{\equiv}CH$), 3.99 (2H, s, C*H*$_2$COO), 4.04 (2H, m, C*H*$_2$NH), 4.27 (2H, q, *J* 7.2, OC*H*$_2$), 9.23 (1H, br s, N*H*); δ_C (50 MHz; $CDCl_3$) 13.1 (*C*H$_3$), 35.7 (*C*H$_2$NH), 45.3 (*C*H$_2$COO), 62.3 (O*C*H$_2$), 71.5 (C*C*H), 77.9 (*C*CH), 165.5 (*C*OO).

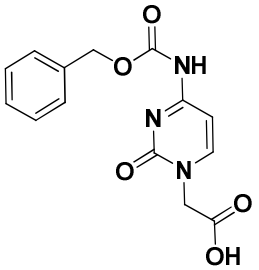
Preparation of *N*^{4}-(benzyloxycarbonyl)cytosine (113)[1]



$C_{12}H_{11}N_3O_3$

Cytosine (**112**; 1.60 g, 14.4 mmol) was suspended in anhydrous pyridine (60 mL) and the mixture was cooled to 0 °C. Benzyl chloroformate (4.11 mL, 28.8 mmol) was added dropwise and the suspension was stirred overnight at room temperature. After this time the solvent was evaporated and water (30 mL) was added. The residue was acidified to pH 1.0 using 4 M hydrochloric acid solution. The crude product was collected by filtration and washed with water. The wet precipitate was suspended in absolute ethanol (15 mL) and heated to reflux for 10 min. The following mixture was then cooled to 10 °C and filtered. The resulting solid collected was washed with diethyl ether (20 mL) and dried *in vacuo* to yield the title compound (**113**; 1.91g, 7.78 mmol; 54%) as a white solid: m.p. > 250 °C (lit.[REF] > 250 °C); (Found C, 58.54; H, 4.22; N, 17.11% $C_{12}H_{11}N_3O_3$ requires C, 58.77; H, 4.52; N, 17.11%); υ_{max}(KBr)/cm^{-1} 3141s, 2959s, 2800s, 1745s, 1689s, 1631s, 1589m, 1514m.

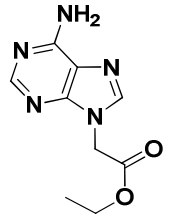
Preparation of N^4-benzyloxycarbonyl cytosin-1-yl acetic acid (107)[1]



$C_{14}H_{13}N_3O_5$

N^4-(Benzyloxycarbonyl)cytosine (**113**; 0.267 g, 1.09 mmol) and anhydrous potassium carbonate (0.151 g, 1.09 mmol) were suspended in dimethylformamide (5 mL). To this heterogeneous mixture was added ethyl bromoacetate (0.13 mL, 1.1 mmol) and the suspension was stirred at room temperature for 18 h. After this time, the solvent was removed under reduced pressure and the resulting residue was treated with 4 M hydrochloric acid solution (1 mL) and stirred for 15 min at 0 °C. The precipitate was collected by filtration and treated with water (4 mL) and 2 M sodium hydroxide solution (2 mL). The resulting solution was stirred for 30 min at room temperature before being cooled down to 0 °C. A white solid precipitated by the addition of 4 M hydrochloric acid solution. This crude product was collected and recrystallised from boiling methanol and washed with diethyl ether to yield the title compound (**107**; 0.10 g, 0.33 mmol; 30%) as a white solid: m.p. 266-274 °C (lit.[1] 266-274 °C); (Found C, 55.39; H, 4.26; N, 13.87% $C_{14}H_{13}N_3O_5$ requires C, 55.45; H, 4.32; N, 13.86%); υ_{max}(KBr)/cm^{-1} 3449br s, 3238m, 3155m, 1747s, 1713s, 1637s; δ_H (200 MHz; DMSO-d_6) 4.52 (2H, s, CH_2COOH), 5.19 (2H, s, OCH_2Ar), 7.02 (1H, d, *J* 7.3, C(5)H), 7.31-7.44 (5H, m, Ar-H), 8.04 (1H, d, *J* 7.3, C(6)H); δ_C (50 MHz; DMSO-d_6) 50.5 (N*C*H$_2$CO), 66.5 (O*C*H$_2$Ar), 93.9 (C(5)), 127.9 (Ar-*C*H), 128.1 (Ar-*C*H), 128.4 (Ar-*C*H), 135.9 (Ar-*C*), 150.3 (C(6)), 153.1 (NH*C*OO), 155.0 (C(2)O), 163.2 (C(4)), 169.3 (*C*OOH).

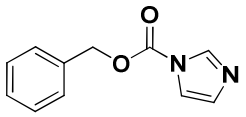
Preparation of ethyl adenin-9-ylacetate (112)[1]



$C_9H_{11}N_5O_2$

Adenine (**114**; 10.0 g, 74.0 mmol) was suspended in anhydrous dimethylformamide (200 mL) and sodium hydride (60% in mineral oil; 3.55 g, 89.0 mmol) was added in small portions. The reaction mixture was stirred for 2 h after which time ethyl bromoacetate (24.7 g, 148 mmol) was added dropwise at room temperature. The mixture was then stirred overnight. After this time, the solvent was evaporated and the resulting oil was shaken with water (100 mL). The crystalline compound formed was collected by filtration, washed with water and recrystallised from boiling methanol to yield the title compound (**115**; 6.64 g, 30.0 mmol; 41%) as white crystals: m.p. 227-228 °C (lit.[1] 227-228 °C); δ_H (200 MHz; DMSO-d_6) 1.14 (3H, t, *J* 7.1, OCH_2CH_3), 4.10 (2H, q, *J* 7.1, OCH_2CH_3), 4.99 (2H, s, NCH_2CO), 7.22 (2H, s, NH_2), 8.04 (2H, s, C(8)H), 8.06 (2H, s, C(2)H); δ_C (50 MHz; DMSO-d_6) 14.4 (OCH_2CH_3), 53.3 (NCH_2CO), 61.8 (OCH_2CH_3), 118.5 (C(5)), 141.7 (C(8)H), 150.1 (C(2)H), 153.0 (C(4)), 156.4 (C(6)), 168.4 (*C*OO).

Preparation of N^1-(benzyloxycarbonyl)imidazole (117)[173, 174]

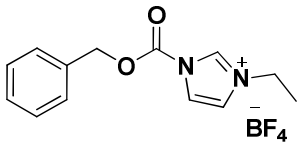


$C_{11}H_{10}N_2O_2$

Benzyl chloroformate (**116**; 12.5 g, 73.2 mmol) was added to a solution of imidazole (10.0 g, 166 mmol) in dry toluene (100 mL) and the mixture was stirred overnight at room temperature. The mixture was then filtered and the filtrate was evaporated to yield an oil which solidified on refrigeration to afford the title compound (**117**; 14.8 g, 73.2 mmol; 100%) as an off-white solid; δ_H (200 MHz; $CDCl_3$) 5.30 (2H, s, CH_2O), 6.93-6.95 (1H, m, Ar-H), 7.18-7.40 (6H, m, Ar-H), 8.02-8.05 (1H, m, Ar-H); δ_C (50 MHz;

$CDCl_3$) 69.8 (*C*H_2O), 117.1 (Ar-*C*H), 128.4 (Ar-*C*H), 128.7 (Ar-*C*H), 129.2 (Ar-*C*H), 130.6 (Ar-*C*H), 133.9 (Ar-*C*), 137.1 (Ar-*C*H), 148.6 (N*C*OO).

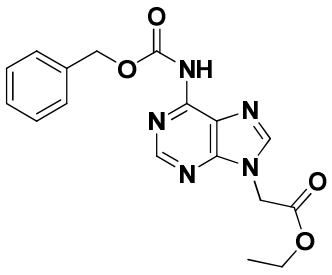
Preparation of 1-(benzyloxycarbonyl)-3-ethylimidazolium tetrafluoroborate (118)[173, 174]



$C_{13}H_{15}BF_4N_2O_2$

Triethyloxonium tetrafluoroborate (2.57 g, 13.5 mmol) was added to N^1-(benzyloxycarbonyl)imidazole (**117**; 2.60 g, 12.9 mmol) in dichloromethane (35 mL) and the mixture was stirred at room temperature for 2 h. The title compound (**118**) was kept in solution without further purification for the subsequent reaction (a small amount of the solution was evaporated to allow characterization); δ_H (200 MHz; $CDCl_3$) 1.35 (3H, t, *J* 7.3, $NCH_2C\underline{H_3}$), 4.16 (2H, q, *J* 7.3, $NC\underline{H_2}CH_3$), 5.34 (2H, s, C*H*$_2$O), 7.16-7.42 (7H, m, Ar-H), 7.54-7.59 (1H, m, Ar-H), 9.10 (1H, s, Ar-H); δ_C (50 MHz; $CDCl_3$) 14.6 (CH_2CH_3), 46.3 (NCH_2), 72.6 (*C*H_2O), 120.1 (Ar-*C*H), 123.3 (Ar-*C*H), 128.9 (Ar-*C*H), 129.3 (Ar-*C*H), 129.6 (Ar-*C*H), 132.6 (Ar-*C*), 136.5 (Ar-*C*H), 145.4 (N*C*OO).

Preparation of ethyl N^6-(benzyloxycarbonyl)adenin-9-ylacetate (119)[1]

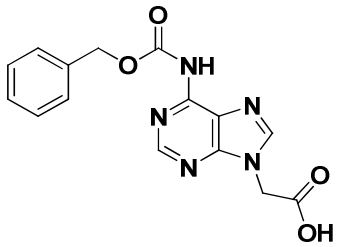


$C_{17}H_{17}N_5O_4$

Ethyl adenin-9-ylacetate (**115**; 3.40 g, 15.4 mmol) was dissolved in dimethylformamide (50 mL) with gentle heating. The cooled solution was then added dropwise to a solution of 1-(benzyloxycarbonyl)-3-ethylimidazolium tetrafluoroborate (**118**; 62.0 mmol) in dichloromethane (35 mL) at 0 °C. The reaction mixture was left to stir at room tempera-

ture overnight. After this time, a saturated sodium bicarbonate solution (100 mL) was added and the resulting mixture was further stirred for 10 min. The organic phase was separated and washed successively with water (100 mL), 1 M potassium hydrogen sulphate solution (2 × 50 mL) and finally brine (50 mL) before being dried over anhydrous magnesium sulphate. Filtration followed by evaporation yielded a colourless oil. This oil was dissolved in dichloromethane (20 mL) and cooled to 0 °C. Light petroleum (50 mL) was then slowly added to cause precipitation of the title compound, which was isolated by filtration (**119**; 4.90 g, 13.8 mmol; 90%) as an off-white solid: m.p. 130-133 °C (lit.[1] 132-135 °C); δ_H (200 MHz; $CDCl_3$) 1.29 (3H, t, *J* 7.1, C*H*$_3$), 4.26 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.95 (2H, s, NC*H*$_2$), 5.29 (2H, s, OC*H*$_2$Ar), 7.33-7.47 (5H, m, 5 × Ar-H), 7.98 (1H, s, C(8)H), 8.68 (1H, s, N*H*), 8.76 (1H, s, C(2)H); δ_C (50 MHz; $CDCl_3$) 14.3 (*C*H$_3$), 44.5 (N*C*H$_2$), 62.7 (O*C*H$_2$CH$_3$), 68.1 (Ar*C*H$_2$O), 123.5 (C(5)), 128.1 (Ar-*C*H), 128.2 (Ar-*C*H), 128.3 (Ar-*C*H), 135.6 (Ar-*C*), 143.3 (C(8)H), 146.8 (C(4)), 149.7 (C(2)H), 151.7 (C(6)), 153.3 (NH*C*OO), 165.4 (*C*OO).

Preparation of N^6-(benzyloxycarbonyl)adenin-9-yl acetic acid (108)[1]

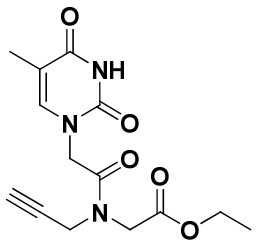


$C_{15}H_{13}N_5O_4$

N^6-(Benzyloxycarbonyl)adenin-9-ylacetate (**119**; 2.50 g, 7.03 mmol) was suspended in methanol (50 mL) at 0 °C and 2 M sodium hydroxide solution (50 mL) was added and the reaction was stirred for 30 min. After this time the mixture was washed with dichloromethane (2 × 50 mL) and the aqueous phase was cooled at 0 °C whereupon the pH was adjusted to 1 using 4 M hydrochloric acid solution. The title compound precipitated as a white solid and was collected by filtration, washed with water and dried *in vacuo* over P_2O_5 (**108**; 2.22 g, 6.78 mmol; 96%); υ_{max}(KBr)/cm^{-1} 3489m, 3086m, 2968m, 2885m, 1737s, 1695s, 1659s, 1592m; δ_H (200 MHz; DMSO-*d*$_6$) 5.12 (2H, s, NC*H*$_2$), 5.29 (2H, s, OC*H*$_2$), 7.35-7.51 (m, 5H, 5 × Ar-H), 8.12 (s, 1H, C(8)H), 8.79 (s, 1H, C(2)H); δ_C (50 MHz; DMSO-*d*$_6$) 45.1 (N*C*H$_2$), 67.3 (O*C*H$_2$), 120.2 (C(5)), 128.0

(Ar-*C*H), 128.2 (Ar-*C*H), 128.5 (Ar-*C*H), 136.2 (Ar-*C*), 145.1 (C(8)H), 148.7 (C(4)), 151.3 (C(2)H), 151.8 (C(6)), 152.3 (NH*C*OO), 168.7 (*C*OOH).

Preparation of *N*-(thymin-1-ylacetyl)-*N*-(prop-2-ynyl)glycine ethyl ester (84a)

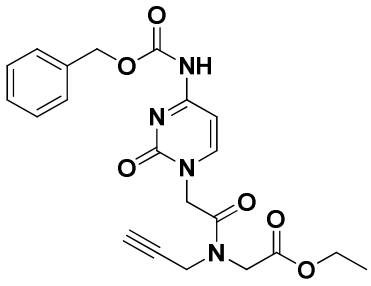


$C_{14}H_{17}N_3O_5$

To a solution of thymine acetic acid (**69**; 0.680 g, 3.70 mmol) in dimethylformamide (10 mL) was added HBTU (1.34 g, 3.53 mmol) in one portion followed by DIPEA (2.0 mL, 12 mmol) and the mixture was stirred for 15 min. In the meantime *N*-(prop-2-ynyl)glycine ethyl ester trifluoroacetate (**106**; 8.86 g, 3.36 mmol) was dissolved in dimethylformamide (10 mL) and then slowly added to the reaction mixture. DIPEA (3.0 mL, 28 mmol) was then added and the homogeneous mixture was stirred overnight at room temperature. After this time, the solvent was removed *in vacuo* and the residue was dissolved in dichloromethane (30 mL). The solution was successively washed with 1 M sodium bicarbonate solution (3 × 15 mL), 1 M potassium hydrogen sulphate solution (2 × 15 mL), water (15 mL) and finally brine (15 mL). The combined organic phases were then dried over magnesium sulphate anhydrous and the solvent was evaporated to an orange foam, which was successively purified by flash column chromatography (ethyl acetate-methanol 95:5) to give the title compound (**84a**; 0.620 g, 2.02 mmol; 60%) as a cream-coloured foam: m.p. 162-163 °C; R_f (ethyl acetate-methanol 95:5) = 0.32; (Found C, 54.77; H, 5.53; N, 13.38% $C_{14}H_{17}N_3O_5$ requires C, 54.72; H, 5.58; N, 13.67%); υ_{max} (film)/cm^{-1} 3238m, 3165w, 2999m, 2832m, 1726s, 1724s, 1682s, 1664s; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.26 and 1.31 (2H, t, *J* 7.1, C*H*$_3$), 1.34 (3H, s, C(5)-C*H*$_3$), 2.30 and 2.48 (1H, t, *J* 2.4, $CH_2C{\equiv}CH$), 4.18 and 4.26 (2H, q, *J* 7.1, OC*H*$_2$), 4.27-4.29 (4H, m, C*H*$_2$NC*H*$_2$), 4.50 and 4.70 (2H, s, NC*H*$_2$CON), 7.01 (1H, s, C(6)H), 9.50 (1H, s, N(3)H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 12.2 (C(5)-*C*H$_3$), 14.0 and 14.1 (*C*H$_3$), 36.3 and 38.0 (HCC*C*H$_2$), 47.7 (*C*H$_2$COO), 48.0 and 48.1 (*C*H$_2$CON), 61.7 and 62.3 (O*C*H$_2$), 73.8 and

74.6 (HC*C*CH$_2$), 76.6 and 77.1 (*C*H), 110.9 (C(5)), 141.1 (C(6)H), 151.3 (C(2)O), 164.5 (C(4)O), 167.1 (CH$_2$*C*ON), 168.7 and 168.9 (*C*OO); *m/z* (ESI) 308 ([M+H]$^+$, 100%), 325 (70), 142 (25); HRMS (ESI) (Found: [M+H]$^+$, 308.1240. $C_{14}H_{18}N_3O_5$ requires *m/z*, 308.1241). An X-Ray crystal structure (Appendix B) was acquired from crystals obtained from diethyl ether and methanol, thereby confirming the structure of the product.

Preparation of *N*-[*N*^{4}-(benzyloxycarbonyl)cytosin-1-ylacetyl]-*N*-(prop-2-ynyl) glycine ethyl ester (84b)

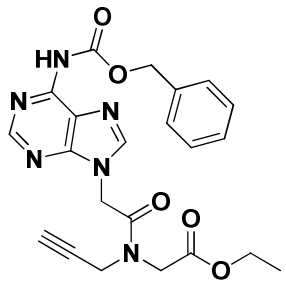


$C_{21}H_{22}N_4O_6$

To a solution of *N*^{4}-benzyloxycarbonyl cytosin-1-yl acetic acid (**107**; 0.32 g, 1.1 mmol) in dimethylformamide (6 mL) was added HBTU (0.36 g, 1.5 mmol) followed by triethylamine (0.28 mL, 2.0 mmol) and the mixture was stirred for 15 min. In the meantime *N*-(prop-2-ynyl)glycine ethyl ester trifluoroacetate (**106**; 0.23 g, 1.02 mmol) was dissolved in dimethylformamide (6 mL) and then slowly added to the reaction mixture. Triethylamine (0.85 mL, 6.1 mmol) was then added and the homogeneous mixture was stirred overnight at room temperature. After this time, dimethylformamide was removed and the residue was dissolved in dichloromethane (10 mL). The solution obtained was successively washed with 1 M sodium bicarbonate solution (3 × 15 mL), 1 M potassium hydrogensulphate solution (2 × 15 mL), water (15 mL) and finally brine (15 mL). The combined organic phase was then dried over magnesium sulphate anhydrous and the solvent was evaporated to an orange foam, which was successively subjected to flash column chromatography (ethyl acetate-methanol 95:5) to give the title compound (**84b**; 0.22 g, 0.52 mmol; 51%) as a white foam: m.p. 130-132 °C; R_f (ethyl acetate) = 0.24; (Found C, 59.02; H, 5.07; N, 13.19% $C_{21}H_{22}N_4O_6$ requires C, 59.15; H, 5.20; N, 13.14%); υ_{max}(KBr)/cm^{-1} 3476m, 3280m, 2984m, 1744s, 1668s, 1629s, 1558s; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.16 and 1.22 (2H, t, *J* 7.1, C*H*$_3$), 2.21 and 2.32 (1H, t, *J* 2.4, C*H*CCH$_2$), 4.07 and 4.20 (2H, q, *J* 7.1, OC*H*$_2$), 4.11 and 4.23 (2H, s, C*H*$_2$COO), 4.17

and 4.18 (2H, s, CHCCH_2), 4.52 and 4.78 (2H, s, CH_2CON), 5.62 (2H, s, OCH_2Ar), 6.96-7.02 (1H, m, C(5)H), 7.26-7.32 (5H, m, 5 × Ar-H), 7.38-7.43 (1H, m, C(6)H), 9.24 (1H, br s, NH-C(4)); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 13.8 (CH_3), 36.3 and 38.2 (HCCCH_2), 47.8 and 48.1 (CH_2COO), 50.8 and 50.9 (CH_2CON), 62.0 and 62.5 (OCH_2), 67.9 (OCH_2Ar), 73.9 and 74.7 (HCCCH$_2$), 76.4 and 76.8 (HCCCH$_2$), 96.3 (C(5)H), 128.4 (Ar-CH), 128.6 (Ar-CH), 128.7 (Ar-CH), 134.9 (Ar-C), 149.8 (C(6)H), 152.2 (NHCO), 157.3 (C(2)O), 163.3 (C(4)), 167.5 and 167.6 (CH$_2$$C$ON), 169.4 and 169.5 ($C$OO); *m/z* (ESI) 427 ([M+H]$^+$, 61%), 319 (3), 85 (100); HRMS (ESI) (Found: [M+H]$^+$, 427.1612. $C_{21}H_{23}N_4O_6$ requires *m/z*, 427.1612).

Preparation of *N*-[N^6-(benzyloxycarbonyl)adenin-9-ylacetyl]-*N*-(prop-2-ynyl) glycine ethyl ester (84c)

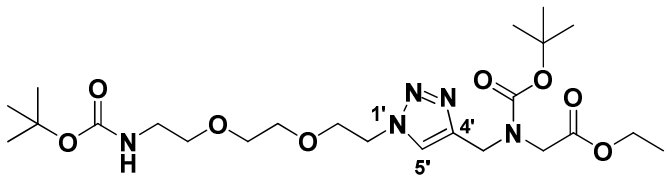


$C_{22}H_{22}N_6O_5$

To a solution of N^6-(benzyloxycarbonyl)adenin-9-yl acetic acid (**108**; 0.332 g, 1.01 mmol) in dimethylformamide (30 mL) was added HBTU (0.402 g, 1.06 mmol) followed by DIPEA (0.19 mL, 1.1 mmol) and the mixture was stirred for 15 min. In the meantime *N*-(prop-2-ynyl)glycine ethyl ester trifluoroacetate (**106**; 0.298 g, 1.17 mmol) was dissolved in dimethylformamide (20 mL) and DIPEA (0.34 mL, 1.93 mmol) was slowly added. This mixture was then added to the solution of **108** and the resulting solution was stirred overnight at room temperature. After this time, dimethylformamide was removed *in vacuo* and the residue was dissolved in dichloromethane (10 mL). The following solution was successively washed with a 1 M sodium bicarbonate solution (3 × 15 mL), 1 M potassium hydrogen sulphate solution (2 × 15 mL), water (15 mL) and finally brine (15 mL). The combined organic phase was then dried over anhydrous magnesium sulphate and the solvent was evaporated to an orange foam, which was successively purified by flash column chromatography (ethyl acetate-methanol 9:1) to give the title

compound (**84c**; 0.164 g, 0.370 mmol; 36 %) as a white solid: m.p. 120-122 °C; R_f (ethyl acetate-methanol 95:5) = 0.39; (Found C, 58.49; H, 4.62; N, 18.86% $C_{22}H_{22}N_6O_5$ requires C, 58.66; H, 4.92; N, 18.66%); υ_{max}(KBr)/cm^{-1} 3237s, 3190s, 3125m, 3029m, 2957m, 1763s, 1749s, 1646s, 1610s; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.24 and 1.32 (2H, t, *J* 7.1, C*H*$_3$), 2.30 and 2.48 (1H, t, *J* 2.4, C*H*CCH$_2$), 4.18 and 4.28 (2H, q, *J* 7.1, OC*H*$_2$), 4.25 and 4.38 (2H, s, C*H*$_2$COO), 4.31 and 4.32 (2H, s, CHCC*H*$_2$), 4.98 and 5.17 (2H, s, C*H*$_2$CON), 5.28 (2H, s, OC*H*$_2$Ar), 7.32-7.42 (5H, m, 5 × Ar-H), 8.02 and 8.05 (1H, s, C(8)H), 8.71 and 8.74 (1H, s, C(2)H), 8.86 (1H, br s, N*H*-C(4)); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.0 and 14.1 (*C*H$_3$), 36.4 and 38.1 (HCC*C*H$_2$), 43.7 and 43.8 (*C*H$_2$CON), 47.5 and 47.9 (*C*H$_2$COO), 61.6 and 62.3 (O*C*H$_2$), 67.1 (O*C*H$_2$Ar), 73.9 and 74.7 (HC*C*CH$_2$), 76.7 and 77.2 (H*C*CCH$_2$), 121.4 (C(5)), 128.4 (Ar-*C*H), 128.6 (Ar-*C*H), 128.7 (Ar-*C*H), 135.4 (Ar-*C*), 143.1 and 143.8 (C(8)H), 149.4 and 149.5 (C(4)), 150.9 and 151.0 (NH*C*OO), 151.4 and 151.5 (C(6)), 152.9 and 153.1 (C(2)H), 165.8 and 165.9 (CH$_2$*C*ON), 167.2 (*C*OO); *m/z* (ESI) 451 ([M+H]$^+$, 100%), 342 (36); HRMS (ESI) (Found: [M+H]$^+$, 451.1725. $C_{22}H_{22}N_6O_5$ requires *m/z*, 451.1724).

Preparation of ethyl 2-{[1-(8-*tert*-butoxycarbonylamino-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl]methyl}*tert*-butoxycarbonylamino acetate (121)

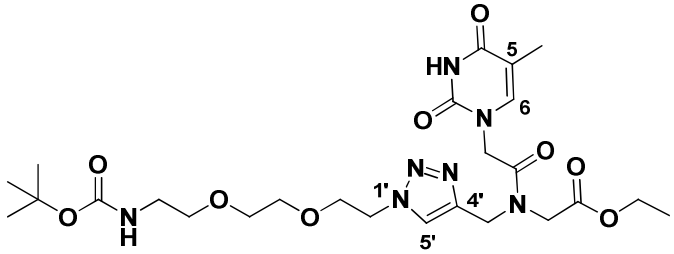


$C_{23}H_{41}N_5O_8$

To the azide **83** (0.089 g, 0.33 mmol) in H_2O/t-BuOH/DCM (9 mL, 1:1:1) was added the monomer **111** (0.088 g, 0.36 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.0081 g, 10 mol %) and 1M sodium ascorbate solution (0.068 mL, 20 mol %). The deep yellow solution was stirred for 24 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated under reduced pressure. The crude residue was subjected to flash column chromatography (ethyl acetate-methanol 9:1) to give the title compound (**121**; 0.127 g, 0.25 mmol, 76 %) as a off-white sticky solid; R_f (ethyl acetate-methanol 9:1) = 0.40; υ_{max}(film)/cm^{-1} 3362m, 2977m, 2933w, 1748s, 1704s,

1514w; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.22 (3H, t, *J* 7.2, C*H*$_3$), 1.38 (9H, s, (C*H*$_3$)$_3$), 1.40 (9H, s, (C*H*$_3$)$_3$), 3.28 (2H, br dt, NHC*H*$_2$), 3.52 (2H, m, NHCH$_2$C*H*$_2$), 3.56 (4H, m, OC*H*$_2$C*H*$_2$O), 3.82 (2H, m, C*H*$_2$CH$_2$N(1')), 3.95 and 3.99 (2H, s, NC*H*$_2$COO), 4.13 (2H, q, *J* 7.2, OC*H*$_2$), 4.44 (2H, m, CH$_2$C*H*$_2$N(1')), 4.55 and 4.57 (2H, s, C(4')C*H*$_2$N), 4.92 and 5.20 (1H, br t, N*H*CH$_2$), 7.56 and 7.75 (1H, s, C(5')H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.0 and 14.1 (*C*H$_3$), 28.2 ((*C*H$_3$)$_3$), 28.3 ((*C*H$_3$)$_3$), 40.3 (NH*C*H$_2$), 42.8 and 43.4 (C(4')*C*H$_2$), 48.1 and 48.8 (*C*H$_2$COO), 50.1 and 50.2 (*C*H$_2$N(1')), 60.8 and 60.9 (O*C*H$_2$), 69.4 and 69.5 (*C*H$_2$CH$_2$N(1')), 70.0 (O*C*H$_2$), 70.2 (O*C*H$_2$), 70.5 (O*C*H$_2$), 80.5 (*C*(CH$_3$)$_3$), 80.6 (*C*(CH$_3$)$_3$), 123.1 and 123.9 (C(5')H), 144.2 and 144.6 (C(4')), 155.2 (*C*ON), 155.9 (*C*ONH), 169.8 (*C*OO); *m/z* (EI) 515 ([M]$^+$, 20%). Structure was verified by NOE and COSY NMR experiments.

Preparation of ethyl 2-(thymin-1-ylacetyl){[1-(8-*tert*-butoxycarbonylamino-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}amino acetate (97a)

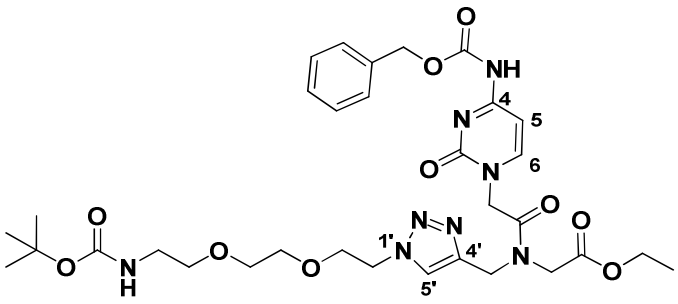


$C_{25}H_{39}N_7O_9$

To the azide **83** (0.090 g, 0.33 mmol) in H_2O/t-BuOH/DCM (9 mL, 1:1:1) was added the monomer **84a** (0.105 g, 0.340 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.0082 g, 10 mol %) and 1 M sodium ascorbate solution (0.068 mL, 21 mol %). The yellow solution was stirred for 48 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrate under reduced pressure. The crude residue was subjected to flash column chromatography (ethyl acetate-methanol 95:5) to give the title compound (**97a**; 0.143 g, 0.250 mmol, 76 %) as a yellow sticky foam. R_f (ethyl acetate-methanol 95:5) = 0.12; υ_{max}(film)/cm^{-1} 3357m, 2978m, 1743s, 1679s, 1678s, 1514w; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.19 and 1.22 (3H, t, *J* 7.1, OCH$_2$C*H*$_3$), 1.42 (9H, s, (C*H*$_3$)$_3$), 1.89

and 1.91 (3H, s, C(5)-CH_3), 3.28-3.30 (2H, m, NHCH_2), 3.50-3.54 (2H, m, NHCH$_2$CH_2), 3.56-3.61 (4H, m, OCH_2CH_2O), 3.78-3.85 (2H, m, CH_2CH$_2$N(1')), 4.14 and 4.28 (2H, s, CH_2COO), 4.16 and 4.18 (2H, q, *J* 7.1, OCH_2CH$_3$), 4.42 and 4.80 (2H, s, CH_2CON), 4.50 and 4.56 (2H, t, *J* 4.9, CH_2N(1')), 4.68 and 4.73 (2H, s, C(4')CH_2), 5.17 and 5.23 (1H, br t, N*H*), 7.02 and 7.05 (1H, s, C(6)H), 7.73 and 7.94 (1H, s, C(5')H), 9.06 and 9.20 (1H, br s, N(3)H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 12.3 (C(5)-*C*H$_3$), 15.2 (OCH$_2$*C*H$_3$), 28.4 ((*C*H$_3$)$_3$), 40.1 (NH*C*H$_2$), 42.6 and 43.4 (C(4')*C*H$_2$), 47.6 and 48.1 (*C*H$_2$COO), 47.8 and 48.9 (*C*H$_2$CON), 50.2 and 50.4 (*C*H$_2$N(1')), 62.1 (O*C*H$_2$CH$_3$), 69.3 (*C*H$_2$CH$_2$N(1')), 69.9 (O*C*H$_2$), 70.2 (O*C*H$_2$), 70.4 (O*C*H$_2$), 79.2 and 79.3 (*C*(CH$_3$)$_3$), 110.8 and 110.8 (C(5)), 123.7 and 124.3 (C(5')H), 140.8 and 140.9 (C(6)H), 144.2 and 144.3 (C(4')), 151.1 (C(4)O), 156.0 (*C*ONH) 164.1 (C(2)O), 167.1 and 167.2 (CH$_2$*C*ON), 168.6 and 168.9 (*C*OO); *m/z* (ESI) 582 ([M+H]$^+$, 100%), 482 (65); HRMS (ESI) (Found: [M+H]$^+$, 582.2881. $C_{25}H_{40}N_7O_9$ requires *m/z*, 582.2882). Structure was determined and verified by NOE NMR experiments.

Preparation of ethyl 2-[*N*^{4}-(benzyloxycarbonyl)cytosin-1-ylacetyl]{[1-(8-*tert*-butoxycarbonylamino-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}amino acetate (97b)

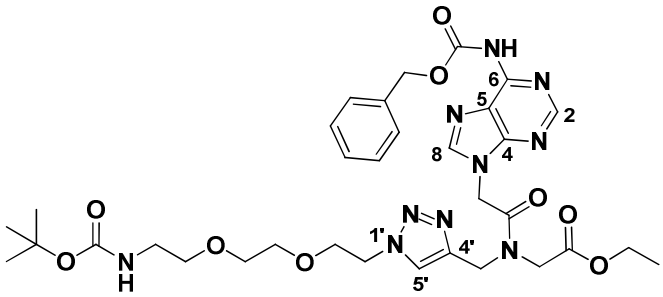


$C_{32}H_{44}N_8O_{10}$

To the azide **83** (0.18 g, 0.68 mmol) in H_2O/t-BuOH/DCM (15 mL, 1:1:1) was added the monomer **84b** (0.29 g, 0.69 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.042 g, 25 mol %) and 1 M sodium ascorbate solution (0.340 mL, 50 mol %). The yellow solution was stirred for 48 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated under reduced pressure. The crude residue was subjected to flash column chromatography (ethyl acetate-methanol 93:7) to yield the title compound (**97b**; 0.32 g, 0.46 mmol, 76 %) as an off-white foam: m.p. 70-73 °C; R_f

(ethyl acetate-methanol 93:7) = 0.26; υ_{max}(film)/cm^{-1} 3262m, 3143m, 2979m, 2931m, 1745s, 1692s, 1668s, 1625s; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.22 and 1.27 (3H, t, *J* 7.1, C*H*$_3$), 1.42 (9H, s, (C*H*$_3$)$_3$), 3.22-3.31 (2H, m, NHC*H*$_2$), 3.47-3.67 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.79-3.88 (2H, m, C*H*$_2$CH$_2$N(1')), 4.09 (2H, br s, C*H*$_2$COO), 4.13 and 4.18 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.39 and 4.88 (2H, s, C*H*$_2$CON), 4.48 and 4.55 (2H, br t, *J* 4.9, C*H*$_2$N(1')), 4.67 and 4.76 (2H, s, C(4')C*H*$_2$), 5.19 (2H, s, OC*H*$_2$Ar), 6.67 and 6.99 (1H, br t, N*H*CH$_2$), 7.20-7.65 (7H, m, 5 × Ar-H, C(5)H and C(6)H), 7.71 and 8.12 (1H, s, C(5')H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 (*C*H$_3$), 28.3 ((*C*H$_3$)$_3$), 40.3 (NH*C*H$_2$), 42.8 and 43.8 (C(4')*C*H$_2$), 47.5 (*C*H$_2$COO), 49.1 (*C*H$_2$CON), 49.7 and 50.3 (*C*H$_2$N(1')), 61.4 (OCH$_2$CH$_3$), 67.9 (O*C*H$_2$Ar), 69.9 (O*C*H$_2$), 70.2 (O*C*H$_2$), 70.4 (O*C*H$_2$), 79.3 (*C*(CH$_3$)$_3$), 95.2 (C(5)H), 123.1 (C(5')H), 128.3 (Ar-*C*H), 128.4 (Ar-*C*H), 128.7 (Ar-*C*H), 134.4 (Ar-*C*), 143.2 (C(4')), 149.9 (C(6)H), 152.1 (NH*C*OO), 153.0 (C(2)O), 155.6 (*C*ONH), 163.2 (C(4)), 167.1 (CH$_2$*C*ON), 168.6 (*C*OO); *m/z* (ESI) 701 ([M+H]$^+$, 100%), 601 (13); HRMS (ESI) (Found: [M+H]$^+$, 701.3243. $C_{32}H_{45}N_8O_{10}$ requires *m/z*, 701.3253).

Preparation of ethyl 2-[*N*^{6}-(benzyloxycarbonyl)adenin-9-ylacetyl]{[1-(8-*tert*-butoxycarbonylamino-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}amino acetate (97c)

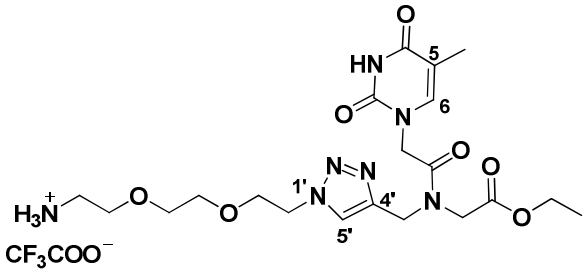


$C_{33}H_{44}N_{10}O_9$

To the azide **83** (0.24 g, 0.87 mmol) in H_2O/t-BuOH/DCM (15 mL, 1:1:1) was added the monomer **84c** (0.470 g, 1.05 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.054 g, 25 mol %) and 1 M sodium ascorbate solution (0.44 mL, 50 mol %). The yellow solution was stirred for 48 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated under reduced pressure. The crude residue was subjected to flash column chromatography (ethyl acetate-methanol 9:1) to yield the title

compound (**97c**; 0.40 g, 0.55 mmol, 64%) as a white foam: m.p. 69-72 °C; R_f (ethyl acetate-methanol 9:1) = 0.20; υ_{max}(film)/cm^{-1} 3362w, 3248w, 3134w, 2988m, 2936m, 2884m, 1747s, 1706s, 1674s, 1612s; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.19 and 1.22 (3H, t, *J* 7.1, OCH_2CH_3), 1.41 (9H, s, $(CH_3)_3$), 3.24-3.32 (2H, m, $NHCH_2$), 3.45 (2H, t, *J* 4.9, $NHCH_2CH_2$), 3.51-3.63 (4H, m, OCH_2CH_2O), 3.79 and 3.86 (2H, t, *J* 4.9, $CH_2CH_2N(1')$), 4.13 and 4.24 (2H, q, *J* 7.1, OCH_2CH_3), 4.18 and 4.42 (2H, s, CH_2COO), 4.50 and 4.60 (2H, t, *J* 4.9, $CH_2N(1')$), 4.71 and 4.84 (2H, s, CH_2CON), 5.01 and 5.42 (2H, s, C(4')CH_2), 5.10 and 5.22 (1H, m, $NHCH_2$), 5.30 (2H, s, OCH_2Ar), 7.33-7.45 (5H, m, Ar-CH), 7.73 and 7.97 (1H, s, C(5')H), 8.12 (1H, br s, C(8)H), 8.72 and 8.76 (1H, s, C(2)H), 8.89 (1H, br s, C(6)-NH); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 and 14.2 (OCH_2CH_3), 28.4 ($(CH_3)_3$), 40.3 ($NHCH_2$), 42.8 and 43.8 (CH_2CON), 43.8 and 44.2 (C(4')CH_2), 47.8 and 49.2 (CH_2COO), 50.2 and 50.5 ($CH_2N(1')$), 61.5 and 62.2 (OCH_2CH_3), 67.7 (OCH_2Ar), 69.1 and 69.2 ($CH_2CH_2N(1')$), 70.1 (OCH_2), 70.4 (OCH_2), 70.6 (OCH_2), 79.5 ($C(CH_3)_3$), 121.4 (C(5)), 123.6 and 124.5 (C(5')H), 128.5 (Ar-CH), 128.6 (Ar-CH), 128.7 (Ar-CH), 135.5 (Ar-C), 141.8 and 142.3 (C(4')), 144.1 (C(8)H), 149.3 (C(4)), 151.1 (CH_2CON), 151.4 (C(6)), 152.8 (C(2)H), 156.0 ($CONH$), 166.3 ($NHCOO$), 168.6 and 168.7 (COO); *m/z* (ESI) 725 ($[M+H]^+$, 100%), 649 (15); HRMS (ESI) (Found: $[M+H]^+$, 725.3358. $C_{33}H_{45}N_{10}O_9$ requires *m/z*, 725.3365).

Preparation of ethyl 2-{[1-(8-amino-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl] methyl}(thymin-1-ylacetyl)amino acetate trifluoroacetic acid (120a)

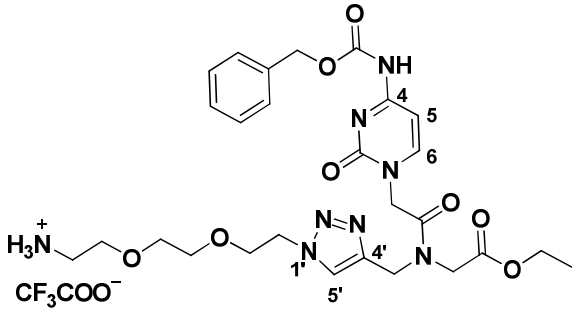


$C_{22}H_{32}F_3N_7O_9$

97a (0.298 g, 0.510 mmol) was dissolved in dichloromethane (5 mL) and trifluoroacetic acid (5 mL) was added. After stirring for 1 h, the solvents were evaporated under reduced pressure to yield a sticky oil, which was used without further purification; δ_H

(400 MHz; D_2O) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.20 (3H, t, *J* 7.1, OCH_2CH_3), 1.90 (3H, s, C(5)-CH_3), 3.28-3.30 (2H, m, $NHCH_2$), 3.50-3.53 (2H, m, $NHCH_2CH_2$), 3.56-3.61 (4H, m, OCH_2CH_2O), 3.78-3.85 (2H, m, $CH_2CH_2N(1')$), 4.14 and 4.28 (2H, s, CH_2COO), 4.19 (2H, q, *J* 7.1, OCH_2CH_3), 4.42 and 4.80 (2H, s, CH_2CON), 4.51-4.53 (2H, m, *J* 4.9, $CH_2N(1')$), 4.68 and 4.73 (2H, s, C(4')CH_2), 7.05 (1H, s, C(6)H), 7.73 and 7.94 (1H, s, C(5')H). Signals corresponding to NH were lost on deuteuration.

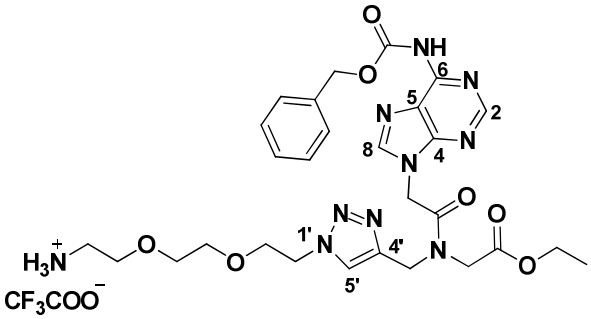
Preparation of ethyl 2-[*N*^{4}-(benzyloxycarbonyl)cytosin-1-ylacetyl]{[1-(8-amino-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}amino acetate trifluoroacetic acid (120b)



$C_{29}H_{37}F_3N_8O_{10}$

97b (0.29 g, 0.51 mmol) was dissolved in dichloromethane (5 mL) and trifluoroacetic acid (5 mL) was added. After stirring for 1 h, the solvents were evaporated under reduced pressure to yield a glassy solid, which was used without further purification; δ_H (200 MHz; D_2O) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.22 and 1.27 (3H, t, *J* 7.1, OCH_2CH_3), 3.24 (2H, br t, $NHCH_2$), 3.47-3.67 (6H, m, $CH_2OCH_2CH_2O$), 3.79-3.88 (2H, m, $CH_2CH_2N(1')$), 4.09 (2H, s, CH_2COO), 4.13 and 4.18 (2H, q, *J* 7.1, OCH_2CH_3), 4.39 and 4.88 (2H, s, CH_2CON), 4.48 and 4.55 (2H, br t, $CH_2N(1')$), 4.67 and 4.76 (2H, s, C(4')CH_2), 5.19 (2H, s, OCH_2Ar), 7.20-7.65 (7H, m, 5 × Ar-CH, C(5)H and C(6)H), 7.71 and 8.12 (s, 1H, C(5')H).

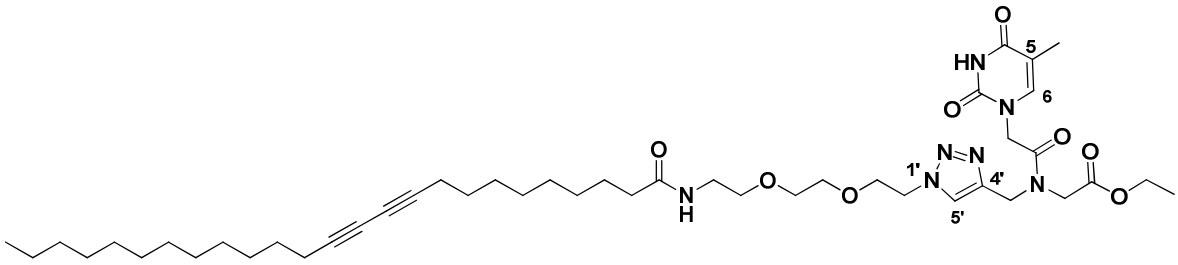
Preparation of ethyl 2-[N^6-(benzyloxycarbonyl)adenin-9-ylacetyl]{[1-(8-amino-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}amino acetate trifluoroacetic acid (120c)



$C_{30}H_{37}F_3N_{10}O_9$

97c (0.37 g, 0.50 mmol) was dissolved in dichloromethane (5 mL) and trifluoroacetic acid (5 mL) was added. After stirring for 1 h, the solvents were evaporated under reduced pressure to yield a glassy solid, which was used without further purification; δ_H (200 MHz; D_2O) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.22 (3H, t, *J* 7.1, C*H*$_3$), 3.24-3.32 (2H, m, NHC*H*$_2$), 3.46-3.63 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.79 and 3.86 (2H, t, *J* 4.9, C*H*$_2$CH$_2$N(1')), 4.25 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.18 and 4.42 (2H, s, NC*H*$_2$COO), 4.53-4.59 (2H, m, *J* 4.9, C*H*$_2$N(1')), 4.71 and 4.84 (2H, s, C*H*$_2$CON), 5.01 and 5.42 (2H, s, C(4')C*H*$_2$), 5.30 (2H, s, OC*H*$_2$Ar), 7.33-7.45 (5H, m, 5 × Ar-CH), 7.73 and 7.97 (1H, s, C(5')H), 8.12 (1H, br s, C(8)H), 8.72 and 8.76 (1H, s, C(2)H). Signals corresponding to NH were lost on deuteration.

Preparation of ethyl 2-(thymin-1-ylacetyl){[1-(8-(pentacosa-10,12-diynamido)-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl]methyl}amino acetate (95a)



$C_{45}H_{71}N_7O_8$

Method A

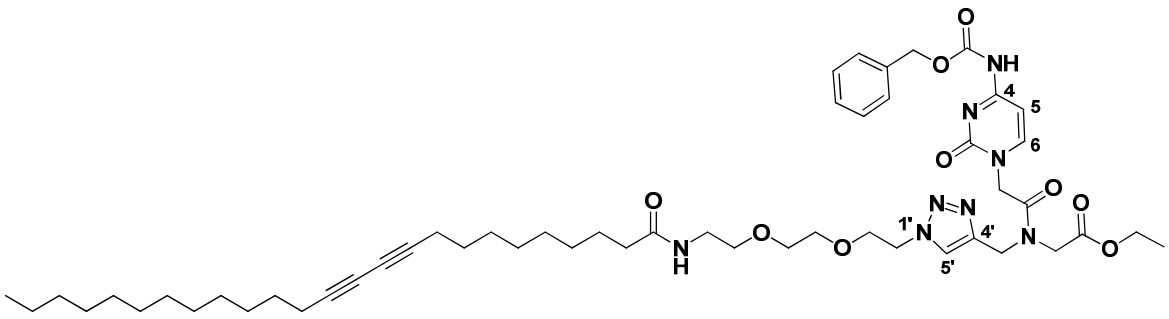
120a (0.304 g, 0.510 mmol) was dissolved in dimethylformamide (6 mL) and triethylamine (0.64 mL, 4.6 mmol) was added. To the stirring solution was added a freshly prepared solution of **78** (0.217 g, 0.570 mmol) in dichloromethane (6 mL). The resulting solution was stirred overnight at room temperature. After this time the solvents were evaporated and the crude material obtained was subjected to flash column chromatography (ethyl acetate-methanol 92:8) to yield the title compound (**95a**; 0.303 g, 0.362 mmol; 71%) as a white solid.

Method B

To the azide **98** (0.10 g, 0.19 mmol) in H_2O/t-BuOH/DCM (12 mL, 1:1:1) was added the propargyl monomer (**84a**; 0.052 g, 0.17 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.0055 g, 13 mol %) and 1 M sodium ascorbate solution (0.034 mL, 20 mol %). The solution was stirred for 48 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated under reduced pressure. The crude residue was subjected to flash column chromatography (ethyl acetate-methanol 92:8) to give the title compound (**95a**; 0.084 g, 0.10 mmol, 59%) as a pale yellow foam: m.p. 90-93 °C; R_f (ethyl acetate-methanol 92:8) = 0.28; υ_{max}(film)/cm^{-1} 3312m, 2926s, 2856s, 1744m, 1711br s, 1675br s, 1539m; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.9, C*H*$_3$), 1.24-1.29 (35H, m, 16 × C*H*$_2$ and OCH$_2$C*H*$_3$), 1.89 and 1.91 (3H, d, $J_{CH3\text{-}C(6)H}$ 1.2, C(5)-C*H*$_3$), 2.15 (2H, t, *J* 7.6, C*H*$_2$CONH), 2.21 (4H, t, *J* 7.1, 2 × (C*H*$_2$C≡C)), 3.38-3.42 (2H, m, NHC*H*$_2$), 3.49-3.51 (2H, m, NHCH$_2$C*H*$_2$), 3.54-3.61 (4H, m, OC*H*$_2$C*H*$_2$O), 3.83 and 3.84 (2H, t, *J* 4.8, C*H*$_2$CH$_2$N(1')), 4.13 and 4.27 (2H, s, C*H*$_2$COO), 4.15 and

4.21 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.39 and 4.76 (2H, s, C*H*$_2$CON), 4.49 and 4.54 (2H, t, *J* 4.8, C*H*$_2$N(1')), 4.66 and 4.73 (2H, s, C(4')C*H*$_2$), 6.41 and 6.47 (1H, br t, N*H*), 7.01 and 7.07 (1H, d, $J_{C(6)H\text{-}CH3}$ 1.2, C(6)H), 7.71 and 7.94 (1H, s, C(5')H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 12.1 (C(5)-*C*H$_3$), 14.1 (OCH$_2$*C*H$_3$ and *C*H$_3$), 19.2 (*C*H$_2$C≡C), 22.6 (*C*H$_2$), 25.7 (*C*H$_2$CH$_2$CONH), 28.2 (*C*H$_2$), 28.3 (*C*H$_2$), 28.8 (*C*H$_2$), 28.9 (*C*H$_2$), 29.0 (*C*H$_2$), 29.1 (*C*H$_2$), 29.2 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 31.9 (*C*H$_2$), 36.6 (*C*H$_2$CONH), 39.1 (NH*C*H$_2$), 42.9 and 43.7 (C(4')*C*H$_2$), 47.7 and 49.0 (*C*H$_2$COO), 47.9 and 48.0 (*C*H$_2$CON), 50.2 and 50.4 (*C*H$_2$N(1')), 61.5 and 62.1 (O*C*H$_2$CH$_3$), 65.2 and 65.3 (*C*≡C-C≡*C*), 69.2 (*C*H$_2$CH$_2$N(1')), 70.0 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.3 (O*C*H$_2$), 77.7 and 77.6 (C≡*C*-*C*≡C), 110.7 and 110.9 (C(5)), 123.7 and 124.2 (C(5')H), 140.8 and 141.0 (C(6)H), 142.4 and 142.8 (C(4')), 151.0 (C(4)O), 164.2 (C(2)O), 167.1 and 167.4 (CH$_2$*C*ON), 168.7 and 168.9 (*C*OO), 173.6 (*C*ONH); *m/z* (ESI) 838 ([M+H]$^+$, 100%); HRMS (ESI) (Found: [M+H]$^+$, 838.5443. $C_{45}H_{72}N_7O_8$ requires *m/z*, 838.5437). Three signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of ethyl 2-(*N*^{4}-(benzyloxycarbonyl)-cytosin-1-ylacetyl){[1-(8-(pentacosa-10,12-diynamido)-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methy} amino acetate (95b)



$C_{52}H_{76}N_8O_9$

Method A

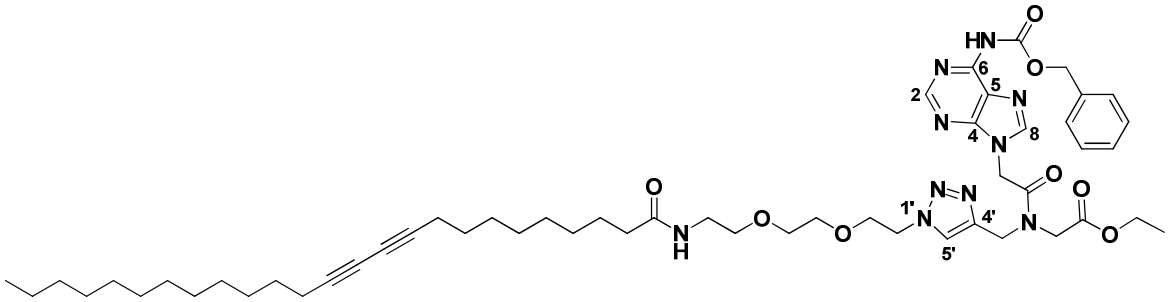
120b (0.25 g, 0.35 mmol) was dissolved in dimethylformamide (5 mL) and triethylamine (0.50 mL, 3.6 mmol) was added. To the stirring solution was added a freshly prepared solution of pentacosa-10,12-diynoyl fluoride (**78**; 0.17 g, 0.44 mmol) in dichloromethane (5 mL). The resulting solution was stirred for 48 h at room temperature.

After this time the solvent was evaporated and the crude material obtained was subjected to flash column chromatography (ethyl acetate-methanol 92:8) yielding the title compound (**95b**; 0.24 g, 0.25 mmol; 70%) as a white sticky foam.

Method B

To the azide **98** (0.24 g, 0.45 mmol) in H_2O/t-BuOH/DCM (12 mL, 1:1:1) was added the monomer **84b** (0.18 g, 0.42 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.011 g, 11 mol %) and 1 M sodium ascorbate solution (0.088 mL, 21 mol %). The solution was stirred for 48 h at room temperature. After this time the solvent was evaporated and the crude residue was subjected to flash column chromatography (ethyl acetate-methanol 95:5) to yield the title compound (**95b**; 0.12 g, 0.13 mmol, 31%) as a colourless sticky foam; R_f (ethyl acetate-methanol 95:5) = 0.18; υ_{max}(film)/cm^{-1} 3300w, 2926s, 2855s, 1746s, 1668br s, 1628s, 1550m; δ_H (200 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.9, C*H*$_3$), 1.19-1.31 (35H, m, 16 × C*H*$_2$ and OCH$_2$C*H*$_3$), 2.12-2.25 (6H, m, C*H*$_2$CONH and 2 × C*H*$_2$C≡C), 3.37-3.41 (2H, m, NHC*H*$_2$), 3.47-3.54 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.82 and 3.86 (2H, t, *J* 4.8, C*H*$_2$CH$_2$N(1')), 4.08 and 4.32 (2H, s, C*H*$_2$COO), 4.16 and 4.20 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.48 and 4.55 (2H, t, *J* 4.8, C*H*$_2$N(1')), 4.55 and 4.88 (2H, s, C*H*$_2$CON), 4.66 and 4.75 (2H, s, C(4')C*H*$_2$), 5.21 (2H, s, OC*H*$_2$Ar), 6.27 and 6.32 (1H, br t, N*H*CH$_2$), 7.25-7.28 (1H, m, C(5)H), 7.35-7.40 (5H, m, 5 × Ar-CH), 7.22-7.65 (1H, m, C(6)H), 7.22 and 8.18 (1H, s, C(5')H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 (OCH$_2$*C*H$_3$ and *C*H$_3$), 19.1 (*C*H$_2$C≡C), 22.7 (*C*H$_2$), 28.3 (*C*H$_2$), 28.7 (*C*H$_2$), 28.8 (*C*H$_2$), 28.9 (*C*H$_2$), 29.1 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.7 (*C*H$_2$), 31.9 (*C*H$_2$), 36.6 (*C*H$_2$CONH), 39.2 (NH*C*H$_2$), 43.1 and 43.9 (C(4')*C*H$_2$), 47.2 and 49.1 (*C*H$_2$COO), 49.5 and 49.6 (*C*H$_2$CON), 50.2 and 50.4 (*C*H$_2$N(1')), 61.5 and 62.1 (OC*H*$_2$CH$_3$), 65.2 and 65.3 (*C*≡C-C≡*C*), 68.3 (OC*H*$_2$Ar), 69.2 (*C*H$_2$CH$_2$N(1')), 69.3 (OC*H*$_2$), 70.0 (OC*H*$_2$), 70.5 (OC*H*$_2$), 77.5 and 77.6 (C≡*C*-*C*≡C), 95.1 (C(5)H), 124.2 and 124.3 (C(5')H), 128.3 (Ar-CH), 128.4 (Ar-CH), 128.7 (Ar-CH), 134.8 (Ar-C), 142.5 and 142.8 (C(4')), 149.6 (C(6)H), 152.1 (NH*C*O), 155.6 (C(2)O), 162.8 (C(4)), 166.7 and 167.2 (CH$_2$*C*ON), 168.6 and 169.1 (*C*OO), 173.4 and 173.6 (*C*ONH); *m/z* (ESI) 957 ([M+H]$^+$, 100%), 869 (3); HRMS (ESI) (Found: [M+H]$^+$, 957.5797. $C_{52}H_{77}N_8O_9$ requires *m/z*, 957.5808). Five signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of ethyl 2-[*N*^{6}-(benzyloxycarbonyl)-adenine-9-ylacetyl]{[1-(8-(pentacosa-10,12-diynamido)-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methy}amino acetate (95c)

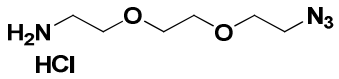


$C_{53}H_{76}N_{10}O_8$

120c (0.37 g, 0.50 mmol) was dissolved in dimethylformamide (8 mL) and triethylamine (0.80 mL, 5.7 mmol) was added. To the stirring solution was added a freshly prepared solution of pentacosa-10,12-diynoyl fluoride (**78**; 0.21 g, 0.55 mmol) in dichloromethane (8 mL). The resulting solution was stirred for 48 h at room temperature. After this time the solvent was evaporated and the crude material obtained was subjected to flash column chromatography (ethyl acetate-methanol 92:8) yielding the title compound (**95c**; 0.36 g, 0.37 mmol; 74%) as a white sticky solid; R_f (ethyl acetate-methanol 92:8) = 0.31; υ_{max}(film)/cm^{-1} 3313m, 2928s, 2861s, 1745m, 1710br s, 1673br s, 1539m; δ_H (200 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.89 (3H, t, *J* 6.9, C*H*$_3$), 1.26-1.57 (35H, m, 16 × C*H*$_2$ and OCH$_2$C*H*$_3$), 2.11-2.19 (2H, m, C*H*$_2$CONH), 2.22-2.28 (4H, m, 2 × C*H*$_2$C≡C), 3.38-3.40 (2H, m, NHC*H*$_2$), 3.43-3.66 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.86 and 3.93 (2H, t, *J* 4.8, C*H*$_2$CH$_2$N(1')), 4.19 and 4.26 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.20 and 4.44 (2H, s, C*H*$_2$COO), 4.52 and 4.61 (2H, t, *J* 4.8, C*H*$_2$N(1')), 4.71 and 4.87 (2H, s, C*H*$_2$CON), 5.05 and 5.43 (2H, s, C(4')C*H*$_2$), 5.33 (2H, s, OC*H*$_2$Ar), 6.09 and 6.15 (1H, br t, N*H*CH$_2$), 7.36-7.49 (5H, m, 5 × Ar-CH), 7.74 and 8.03 (1H, s, C(5')H), 8.11 and 8.13 (1H, s, C(8)H), 8.42 (1H, br s, C(6)-N*H*), 8.74 and 8.78 (1H, s, C(2)H); δ_C (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 (OCH$_2$*C*H$_3$ and *C*H$_3$), 19.2 (*C*H$_2$C≡C), 22.7 (*C*H$_2$), 28.3 (*C*H$_2$), 28.4 (*C*H$_2$), 28.8 (*C*H$_2$), 28.9 (*C*H$_2$), 29.1 (*C*H$_2$), 29.2 (*C*H$_2$), 29.3 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.7 (*C*H$_2$), 31.9 (*C*H$_2$), 36.6 (*C*H$_2$CONH), 39.1 (NH*C*H$_2$), 42.9 and 43.8 (*C*H$_2$CON), 43.9 and 44.2 (C(4')*C*H$_2$), 47.9 and 49.3 (*C*H$_2$COO), 50.2 and 50.5 (*C*H$_2$N(1')), 61.6

and 62.3 (O*C*$H_2$$CH_3$), 65.2 and 65.8 (*C*≡C-C≡*C*), 67.8 (O*C*H_2Ar), 69.2 and 69.3 (*C*$H_2$$CH_2$N(1')), 69.9 (O*C*$H_2$), 70.0 (O*C*$H_2$), 70.4 (O*C*$H_2$), 77.2 and 77.5 (C≡*C*-*C*≡C), 121.2 (C(5)), 123.8 and 124.5 (C(5')H), 128.5 (Ar-CH), 128.6 (Ar-CH), 128.7 (Ar-CH), 135.4 (Ar-C), 142.3 and 142.5 (C(4')), 143.8 and 144.0 (C(8)H), 149.3 (C(4)), 150.8 (NH*C*OO), 151.5 (C(6)), 152.8 (C(2)H), 166.4 (CH_2*C*ON), 168.6 and 168.7 (*C*OO), 173.3 (*C*ONH); *m/z* (ESI) 981 ($[M+H]^+$, 100%), 905 (12), 752 (10); HRMS (ESI) (Found: $[M+H]^+$, 981.5919. $C_{53}H_{77}N_{10}O_8$ requires *m/z*, 981.5920). Five signals corresponding to *C*H_2 from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

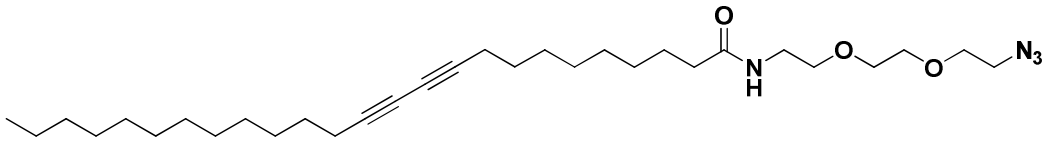
Preparation of 8-amino-3,6-dioxaoctan-1-azide hydrochloride (122)



$C_6H_{15}N_4O_2Cl$

Compound **83** (0.776 g, 2.83 mmol) was dissolved in ethyl acetate (15 mL) and 6 M hydrochloric acid (15 mL) was added. The mixture was stirred for 1 h and after this time the solvents were evaporated to yield the title compound (**122**; 0.589 g, 2.80 mmol; 99%) as a sticky orange oil; υ_{max}(film)/cm^{-1} 3400m, 2927m, 2874m, 2112s (N_3 str), 1654w; δ_H (400 MHz; MeOH-*d*$_4$) 3.12 (2H, br t, *J* 5.3, NHC*H*$_2$), 3.38 (2H, t, *J* 5.3, C*H*$_2$$N_3$), 3.63-3.71 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.75 (2H, t, *J* 5.3, C*H*$_2$$CH_2N_3$); δ_C (50 MHz; MeOH-*d*$_4$) 41.0 (NH*C*H_2), 52.1 (*C*$H_2$$N_3$), 68.2 (*C*$H_2$$CH_2N_3$), 71.4 (O*C*$H_2$), 71.7 (O*C*$H_2$), 71.8 (O*C*$H_2$). Signal corresponding to N*H*$_2$ protons were exchanged and were not seen.

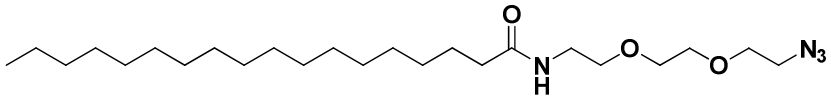
Preparation of *N*-(8-azido-3,6-dioxaoctanyl)pentacosa-10,12-diynamide (98)



$C_{31}H_{54}N_4O_3$

To a solution of **122** (0.589 g, 2.80 mmol) in dimethylformamide (10 mL) was added triethylamine (1.95 mL, 13.9 mmol) followed by a solution of pentacosa-10,12-diynoyl fluoride (**78**; 1.15 g, 3.05 mmol) in dichloromethane (10 mL). The mixture was stirred for 20 h at room temperature. After this time the solvent was removed under reduced pressure and the residue was dissolved in dichloromethane (20 mL). This was then washed with 1 M sodium bicarbonate (2 × 10 mL), 1 M potassium hydrogen sulphate (2 × 10 mL), water (10 mL) and brine (10 mL). The organic layer was dried with magnesium sulphate and the solvent removed *in vacuo*. The resulting crude solid was subjected to flash column chromatography (ethyl acetate) to yield the title compound (**98**; 1.30 g, 2.96 mmol, 80%) as a white-off blue solid: m.p. 70-72 °C; R_f (ethyl acetate) = 0.34; υ_{max}(film)/cm^{-1} 3300s, 3084w, 2919s, 2848s, 2108s (N_3 str), 1643s and 1557m; δ_H (400 MHz, $CDCl_3$) 0.88 (3H, t, *J* 6.3, C*H*$_3$CH$_2$), 1.12-1.30 (32H, m, 16 × C*H*$_2$), 2.15 (2H, t, *J* 7.6, C*H*$_2$CONH), 2.24 (4H, t, *J* 6.7, 2 × C*H*$_2$C≡C), 3.38 (2H, t, *J* 5.0, CH$_2$C*H*$_2$N$_3$), 3.44 (2H, br dt, NHC*H*$_2$CH$_2$O), 3.53 (2H, t, *J* 5.0, NHCH$_2$C*H*$_2$O), 3.60-3.66 (4H, m, OC*H*$_2$C*H*$_2$O), 3.66 (2H, t, *J* 5.0, OC*H*$_2$CH$_2$N$_3$), 5.55 (1H, br s, N*H*); δ_C (100 MHz, $CDCl_3$) 14.1 (*C*H$_3$), 19.1 (*C*H$_2$C≡C), 22.6 (*C*H$_2$), 28.2 (*C*H$_2$), 28.3 (*C*H$_2$), 28.7 (*C*H$_2$), 28.8 (*C*H$_2$), 28.9 (*C*H$_2$), 29.5 (*C*H$_2$), 29.1 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.6 (*C*H$_2$), 31.9 (*C*H$_2$), 36.6 (*C*H$_2$CO), 39.0 (NH*C*H$_2$), 50.6 (*C*H$_2$N$_3$), 65.1 and 65.2 (C≡*C*-*C*≡C), 70.0 (NHCH$_2$*C*H$_2$O), 70.1 (O*C*H$_2$CH$_2$N$_3$), 70.2 (O*C*H$_2$), 70.5 (O*C*H$_2$), 77.4 and 77.5 (*C*≡C-C≡*C*), 173.1 (*C*O); *m/z* (CI) 531 ([M+H]$^+$, 20%), 88 (61), 86 (100), 72 (65), 58 (98); HRMS (ESI) (Found: [M+H]$^+$, 531.4260. $C_{31}H_{55}N_4O_3$ requires *m/z*, 531.4269). Three signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

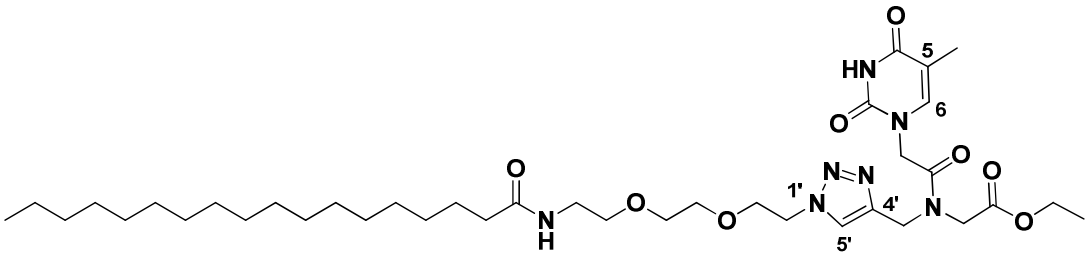
Preparation of *N*-(8-azido-3,6-dioxaoctanyl)stearamide (99)



$C_{24}H_{48}N_4O_3$

To a solution of **122** (0.779 g, 3.70 mmol) in dimethylformamide (20 mL) was added triethylamine (2.60 mL; 18.5 mmol) followed by a solution of stearoyl fluoride (**79**; 1.19 g, 4.15 mmol) in dichloromethane (20 mL). The mixture was stirred for 20 h at room temperature. After this time the solvent was removed under reduced pressure and the residue was dissolved in dichloromethane (40 mL). This was then washed with 1 M sodium bicarbonate (2 × 30 mL), 1 M potassium hydrogen sulphate (2 × 30 mL), water (30 mL) and brine (30 mL). The organic layer was dried over magnesium sulphate and the solvent removed under reduced pressure. The resulting yellow solid was subjected to flash column chromatography (ethyl acetate) to yield the title compound (**99**; 1.30 g, 2.96 mmol, 80%) as a white solid: m.p. 65-66 °C; R_f (ethyl acetate) = 0.25; (Found C, 65.43; H, 11.13; N, 12.58% $C_{24}H_{48}N_4O_3$ requires C, 65.41; H, 10.98; N, 12.71%); υ_{max}(film)/cm^{-1} 3307s, 3083w, 2918s, 2850s, 2107s (N_3 str), 1641s and 1553m; δ_H (400 MHz, $CDCl_3$) 0.88 (3H, t, *J* 6.3, C*H*$_3$), 1.26 (28H, s, 14 × C*H*$_2$), 1.61 (2H, m, C*H*$_2$CH$_2$CO), 2.14 (2H, t, *J* 7.6, C*H*$_2$CONH), 3.38 (2H, t, *J* 5.0, CH$_2$C*H*$_2$N$_3$), 3.44 (2H, br dt, NHC*H*$_2$CH$_2$O), 3.54 (2H, t, *J* 5.0, NHCH$_2$C*H*$_2$O), 3.60-3.66 (4H, m, OC*H*$_2$C*H*$_2$O), 3.67 (2H, t, *J* 5.0, OC*H*$_2$CH$_2$N$_3$), 5.55 (1H, br s, N*H*); δ_C (100 MHz, $CDCl_3$) 14.0 (*C*H$_3$), 22.6 (*C*H$_2$), 25.7 (*C*H$_2$CH$_2$CO), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.6 (*C*H$_2$), 29.7 (*C*H$_2$), 31.9 (*C*H$_2$), 36.7 (*C*H$_2$CO), 39.0 (NH*C*H$_2$), 50.6 (*C*H$_2$N$_3$), 70.0 (NHCH$_2$*C*H$_2$O), 70.1 (O*C*H$_2$CH$_2$N$_3$), 70.2 (O*C*H$_2$), 70.5 (O*C*H$_2$), 173.2 (*C*O); *m/z* (CI) 441 ([M+H]$^+$, 3%), 415 (5), 86 (100), 72 (99); HRMS (ESI) (Found: [M+H]$^+$, 441.3796. $C_{24}H_{49}N_4O_3$ requires *m/z*, 441.3799). Six signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of ethyl 2-(thymin-1-ylacetyl)((1-(8-stearamido-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl)methyl)aminoacetate (96a)

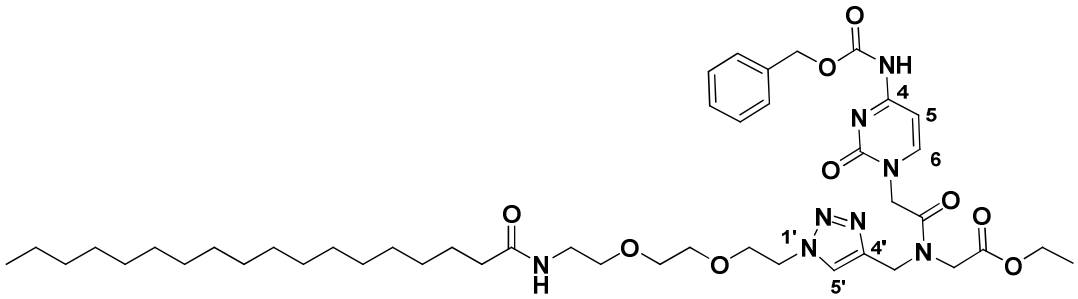


$C_{38}H_{65}N_7O_8$

To the azide **99** (0.162 g, 0.340 mmol) in H_2O/t-BuOH/DCM (15 mL, 1:1:1) was added the monomer **84a** (0.106 mg, 0.370 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.023 g, 27 mol %) and 1 M sodium ascorbate solution (0.185 mL, 54 mol %). The yellow solution was stirred for 48 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated. The crude residue was subjected to flash column chromatography (ethyl acetate-methanol 9:1) to give the title compound (**96a**; 0.187 g, 0.250 mmol, 74%) as a yellow foam: m.p. 66-67 °C; R_f (ethyl acetate-methanol 9:1) = 0.17; υ_{max}(film)/cm^{-1} 3309m, 3064m, 2924s, 2853s, 1744s, 1716br s, 1674br s, 1545m; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.3, C*H*$_3$), 1.26 (31H, s, 14 × C*H*$_2$ and OCH$_2$C*H*$_3$), 1.53-1.58 (2H, m, C*H*$_2$CH$_2$CONH), 1.88 and 1.91 (3H, s, C(5)-C*H*$_3$), 2.15 (2H, t, *J* 7.6, C*H*$_2$CONH), 3.38-3.40 (2H, m, NHC*H*$_2$), 3.49-3.54 (2H, m, NHCH$_2$C*H*$_2$), 3.57-3.60 (4H, m, OC*H*$_2$C*H*$_2$O), 3.79-3.86 (2H, m, C*H*$_2$CH$_2$N(1')), 4.15 and 4.27 (2H, s, C*H*$_2$COO), 4.16 and 4.21 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.39 and 4.75 (2H, s, C*H*$_2$CON), 4.50 and 4.54 (2H, t, *J* 4.9, C*H*$_2$N(1')), 4.68 and 4.72 (2H, s, C(4')C*H*$_2$), 6.31 and 6.32 (1H, br t, N*H*), 7.01 and 7.07 (1H, s, C(6)H), 7.71 and 7.94 (1H, s, C(5')H), 8.98 and 9.23 (1H, br s, N(3)H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 12.3 (C(5)-*C*H$_3$), 14.1 (*C*H$_3$CH$_2$ and OCH$_2$*C*H$_3$), 22.7 (*C*H$_2$), 25.8 (*C*H$_2$CH$_2$CONH), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.7 (*C*H$_2$), 31.9 (*C*H$_2$), 36.6 (*C*H$_2$CONH), 39.1 (NH*C*H$_2$), 42.8 and 43.8 (C(4')*C*H$_2$), 47.6 and 49.1 (*C*H$_2$COO), 47.9 and 48.0 (*C*H$_2$CON), 50.2 and 50.4 (*C*H$_2$N(1')), 61.5 and 62.1 (O*C*H$_2$CH$_3$), 69.3 (*C*H$_2$CH$_2$N(1')), 70.0 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.3 (O*C*H$_2$), 110.8 and 110.9 (C(5)), 123.8 and 124.2 (C(5')H), 140.7 and 141.0 (C(6)H), 142.4 and 142.8 (C(4')), 151.1 (C(4)O), 164.1 (C(2)O), 167.1 and 167.3

(CH_2CON), 168.6 and 168.9 (*C*OO), 173.6 (*C*ONH); *m/z* (CI) 748 ($[M+H]^+$, 21%), 163 (30), 146 (60), 77 (100); HRMS (ESI) (Found: $[M+H]^+$, 748.4977. $C_{38}H_{66}N_7O_8$ requires *m/z*, 748.4967). Seven signals corresponding to CH_2 from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of ethyl 2-[*N*^{4}-(benzyloxycarbonyl)cytosin-1-ylacetyl]{[1-(8-stearamido-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl]methyl}aminoacetate (96b)



$C_{45}H_{70}N_8O_9$

Method A

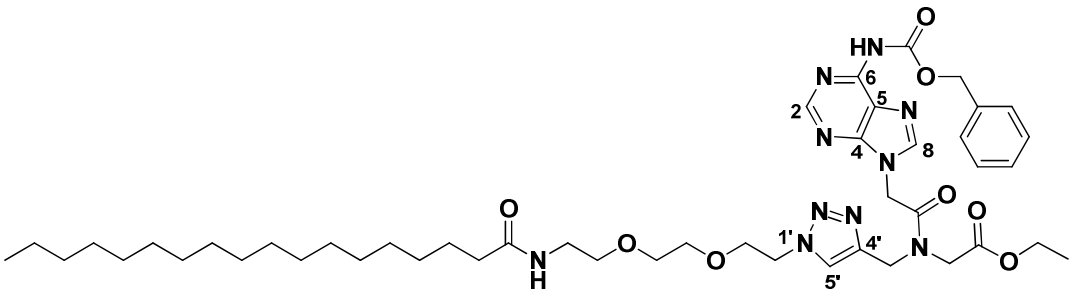
To the azide **99** (0.24 g, 0.54 mmol) in H_2O/t-BuOH/DCM (12 mL, 1:1:1) was added the monomer **84b** (0.23 g, 0.53 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.035 g, 26 mol %) and 1 M sodium ascorbate solution (0.275 mL, 52 mol %). The solution was stirred for 48 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated. The crude residue was subjected to flash column chromatography (ethyl acetate - methanol 93:7) to give the title compound (**96b**; 0.29 g, 0.33 mmol, 62%) as a pale yellow foam.

Method B

107 (0.046 g, 0.15 mmol) was dissolved in dimethylformamide (4 mL) and HBTU (0.056 g, 0.14 mmol) followed by triethylamine (0.022 mL, 0.16 mmol) was added. The mixture was stirred for 15 min at room temperature and to this solution was added a solution of **124** (0.11 g, 0.16 mmol) and triethylamine (0.147 mL, 1.06 mmol) in dimethylformamide (4 mL). The mixture was stirred for 20 h at room temperature and after this time the solvent was evaporated and the crude residue subjected to flash column chromatography (ethyl acetate - methanol 95:5) to isolate the title compound (**96b**; 0.069 g, 0.080 mmol; 53%) as pale yellow powder: m.p. 58-61 °C; R_f (ethyl acetate-methanol 95:5) = 0.18; υ_{max}(film)/cm^{-1} 3298m, 2924s, 2854s, 1746s, 1667br s, 1628m,

1555m; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.3, C*H*$_3$CH$_2$), 1.22 (28H, s, 14 × C*H*$_2$), 1.27 (3H, t, *J* 7.1, OCH$_2$C*H*$_3$), 1.55-1.57 (2H, m, C*H*$_2$CH$_2$CONH), 2.12-2.16 (2H, m, C*H*$_2$CONH), 3.38-3.40 (2H, m, NHC*H*$_2$), 3.49-3.53 (2H, m, NHCH$_2$C*H*$_2$), 3.55-3.58 (4H, m, OC*H*$_2$C*H*$_2$O), 3.82 and 3.86 (2H, t, *J* 4.8, C*H*$_2$CH$_2$N(1')), 4.10 and 4.35 (2H, s, C*H*$_2$COO), 4.15 and 4.21 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.48 and 4.52 (2H, t, *J* 4.8, C*H*$_2$N(1')), 4.53 and 4.89 (2H, s, C*H*$_2$CON), 4.66 and 4.78 (2H, s, C(4')C*H*$_2$), 5.21 (2H, s, OC$\underline{H_2}$Ar), 6.66-6.75 (1H, m, N*H*CH$_2$), 7.22-7.24 (1H, m, C(5)H), 7.35-7.40 (5H, m, 5 × Ar-CH), 7.48-7.52 (1H, m, C(6)H), 7.72 and 8.15 (1H, s, C(5')H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.0 (*C*H$_3$ and OCH$_2$*C*H$_3$), 22.6 (*C*H$_2$), 25.7 (*C*H$_2$CH$_2$CONH), 29.2 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.7 (*C*H$_2$), 31.8 (*C*H$_2$), 36.6 (*C*H$_2$CONH), 39.1 (NH*C*H$_2$), 42.8 and 43.9 (C(4')*C*H$_2$), 47.4 and 49.1 (*C*H$_2$COO), 49.6 and 50.1 (*C*H$_2$CON), 49.8 and 50.3 (*C*H$_2$N(1')), 61.4 and 62.0 (O*C*H$_2$CH$_3$), 67.9 (O*C*H$_2$Ar), 69.2 (*C*H$_2$CH$_2$N(1')), 70.0 (O*C*H$_2$), 70.4 (O*C*H$_2$), 70.5 (O*C*H$_2$), 95.0 (C(5)H), 124.2 and 124.3 (C(5')H), 128.2 (Ar-CH), 128.3 (Ar-CH), 128.6 (Ar-CH), 134.9 (Ar-C), 142.3 and 142.5 (C(4')), 149.6 (C(6)H), 152.2 (NH*C*OO), 157.7 (C(2)O), 162.9 (C(4)), 166.7 and 167.2 (*C*H$_2$CON), 168.6 and 169.0 (*C*OO), 173.4 and 173.5 (*C*ONH); *m/z* (ESI) 701 ([M+H]$^+$, 100%), 601 (13); HRMS (ESI) (Found: [M+H]$^+$, 867.5326. $C_{45}H_{71}N_8O_9$ requires *m/z*, 867.5339). Six signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of ethyl 2-[*N*^{6}-(benzyloxycarbonyl)adenin-9-ylacetyl]{[1-(8-stearamido-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl]methyl}aminoacetate (96c)

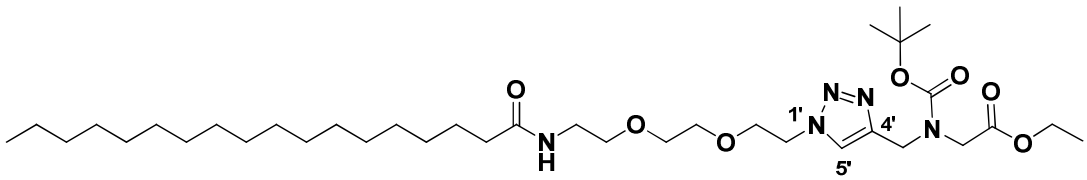


$C_{46}H_{70}N_{10}O_8$

To the azide **99** (0.093 g, 0.21 mmol) in H_2O/t-BuOH/DCM (6 mL, 1:1:1) was added the monomer **84c** (0.091 g, 0.20 mmol), followed by $CuSO_4.5H_2O$ (0.013 g, 26 mol %)

and 1 M sodium ascorbate (0.10 mL, 52 mol %). The solution was stirred for 48 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated. The crude residue was subjected to flash column chromatography (ethyl acetate - methanol 92:8) to give the title compound (**96c**; 0.10 g, 0.11 mmol, 55%) as a white sticky foam; R_f (ethyl acetate-methanol 92:8) = 0.13; υ_{max}(film)/cm^{-1} 3289w, 2916s, 2854s, 1742s, 1664s, 1618s; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.3, C*H*$_3$CH$_2$), 1.16-1.31 (31H, m, 14 × C*H*$_2$ and OCH$_2$C*H*$_3$), 1.54-1.56 (2H, m, C*H*$_2$CH$_2$CONH), 2.08-2.12 (2H, m, C*H*$_2$CONH), 3.30-3.35 (2H, m, NHC*H*$_2$), 3.37-3.42 (2H, m, NHCH$_2$C*H*$_2$), 3.45-3.58 (4H, m, OC*H*$_2$C*H*$_2$O), 3.79 and 3.86 (2H, t, *J* 4.8, C*H*$_2$CH$_2$N(1')), 4.13 and 4.22 (2H, q, *J* 7.1, OC*H*$_2$CH$_3$), 4.14 and 4.38 (2H, s, C*H*$_2$COO), 4.48 and 4.55 (2H, t, *J* 4.8, C*H*$_2$N(1')), 4.56 and 4.81 (2H, s, C*H*$_2$CON), 4.97 and 5.38 (2H, s, C(4')C*H*$_2$), 5.26 (2H, s, OC*H*$_2$Ar), 6.05 and 6.11 (1H, br t, N*H*CH$_2$), 7.31-7.43 (5H, m, 5 × Ar-CH), 7.58 and 7.97 (1H, s, C(5')H), 8.03 and 8.06 (1H, s, C(8)H), 8.66 (1H, br s, C(6)-N*H*), 8.68 and 8.72 (1H, s, C(2)H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 (*C*H$_3$ and OCH$_2$*C*H$_3$), 22.6 (*C*H$_2$), 25.7 (*C*H$_2$CH$_2$CONH), 29.2 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.7 (*C*H$_2$), 31.9 (*C*H$_2$), 36.6 (*C*H$_2$CON), 39.1 (NH*C*H$_2$), 42.8 and 43.5 (*C*H$_2$CON), 43.6 and 44.0 (C(4')*C*H$_2$), 47.8 and 49.2 (*C*H$_2$COO), 50.2 and 50.5 (*C*H$_2$N(1')), 61.5 and 62.2 (O*C*H$_2$CH$_3$), 67.7 (O*C*H$_2$Ar), 69.1 and 69.2 (*C*H$_2$CH$_2$N(1')), 69.9 (O*C*H$_2$), 70.4 (O*C*H$_2$), 70.5 (O*C*H$_2$), 121.4 (C(5)), 123.8 and 124.3 (C(5')H), 128.5 (Ar-CH), 128.6 (Ar-CH), 128.7 (Ar-CH), 135.4 (Ar-C), 142.0 and 142.2 (C(4')), 143.9 (C(8)H), 149.3 (C(4)), 151.1 (NH*C*OO), 151.4 (C(6)), 152.8 (C(2)H), 166.3 and 166.4 (*C*ONH), 168.6 and 168.7 (*C*OO), 173.3 (*C*ONH); *m/z* (ESI) 891 ([M+H]$^+$, 100%), 815 (12), 427 (10), 217 (8); HRMS (ESI) (Found: [M+H]$^+$, 891.5447. $C_{46}H_{71}N_{10}O_8$ requires *m/z*, 891.5451). Six signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

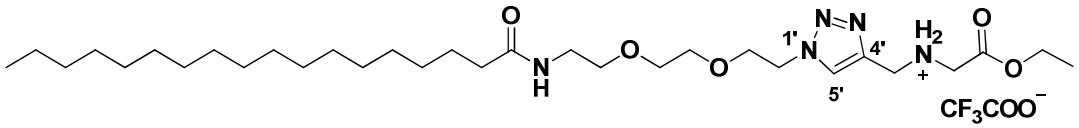
Preparation of ethyl 2-{[1-(8-stearamido-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}*tert*-butoxycarbonylaminoacetate (123)



$C_{36}H_{67}N_5O_7$

To the azide **99** (0.10 g, 0.23 mmol) in H_2O/t-BuOH/DCM (12 mL, 1:1:1) was added the monomer **111** (0.070 g, 0.29 mmol), followed by $CuSO_4 \cdot 5H_2O$ (0.0063 g, 11 mol %) and 1 M sodium ascorbate solution (0.048 mL, 21 mol %). The solution was stirred for 24 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrate under reduced pressure. The crude residue was subjected to flash column chromatography (ethyl acetate-light petroleum ether 9:1) to give the title compound (**123**; 0.14 g, 0.21 mmol, 89 %) as a sticky solid; R_f (ethyl acetate-light petroleum 9:1) = 0.18; (Found C, 63.15; H, 9.78; N, 10.14% $C_{36}H_{67}N_5O_7$ requires C, 63.40; H, 9.90; N, 10.27%); υ_{max}(film)/cm^{-1} 3304m, 2917s, 2850s, 1749m, 1699br s, 1640s, 1556m; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.3, C*H*$_3$), 1.26 (31H, s, 14 × C*H*$_2$ and OCH$_2$C*H*$_3$) 1.35 and 1.40 (9H, s, (C*H*$_3$)$_3$), 1.57 (2H, m, C*H*$_2$CH$_2$CONH), 2.15 (2H, m, C*H*$_2$CONH), 3.36-3.39 (2H, m, NHC*H*$_2$CH$_2$O), 3.49-3.51 (2H, m, NHCH$_2$C*H*$_2$), 3.51-3.58 (4H, m, OC*H*$_2$C*H*$_2$O), 3.80-3.83 (2H, m, C*H*$_2$CH$_2$N(1')), 3.90 and 3.97 (2H, s, C*H*$_2$COO), 4.13 (2H, q, *J* 7.2, OC*H*$_2$CH$_3$), 4.46-4.48 (2H, m, C*H*$_2$N(1')), 4.50 and 4.52 (2H, s, C(4')C*H*$_2$), 6.01 and 6.28 (1H, br t, N*H*), 7.61 and 7.78 (1H, s, C(5')H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 (OCH$_2$*C*H$_3$), 14.1 (*C*H$_3$CH$_2$), 22.6 (*C*H$_2$), 25.7 (*C*H$_2$CH$_2$CONH), 28.1 ((*C*H$_3$)$_3$), 29.1 (*C*H$_2$), 29.2 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 31.8 (*C*H$_2$), 36.6 (*C*H$_2$CONH), 39.2 (NH*C*H$_2$), 42.8 and 43.2 (C(4')*C*H$_2$), 48.0 and 48.9 (*C*H$_2$COO), 50.0 (*C*H$_2$N(1')), 60.9 (O*C*H$_2$CH$_3$), 69.9 (O*C*H$_2$), 70.1 (O*C*H$_2$), 70.3 (O*C*H$_2$), 80.5 (*C*(CH$_3$)$_3$), 123.1 and 123.5 (C(5')H), 144.2 and 144.6 (C(4')), 155.2 (*C*ON), 169.8 (*C*OO), 173.2 (*C*ONH); *m/z* (ESI) 682 ([M+H]$^+$, 100%), 654 (10), 582 (5); HRMS (ESI) (Found: [M+H]$^+$, 682.5102. $C_{36}H_{68}N_5O_7$ requires *m/z*, 682.5113). Six signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

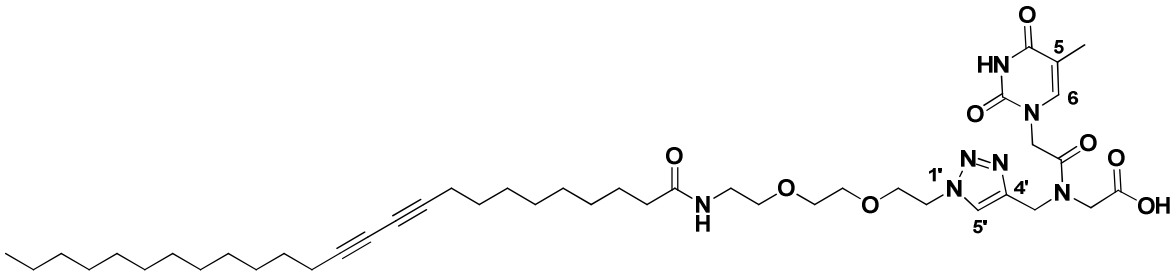
Preparation of ethyl 2-{[1-(8-stearamido-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}aminoacetate trifluoroacetic acid (124)



$C_{33}H_{60}F_3N_5O_7$

Compound **123** (0.138 g, 0.20 mmol) was dissolved with DCM (5 mL) and trifluoroacetic acid (5 mL) was added. The mixture was stirred for 1 h and after this time the solvents were evaporated to yield the title compound (**124;** 0.132 g, 0.19 mmol, 95%) as colourless sticky solid, which was used without further purification; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.3, C*H*$_3$), 1.16-1.37 (31H, m, 14 × C*H*$_2$ and OCH$_2$C*H*$_3$), 1.49-1.60 (2H, m, C*H*$_2$CH$_2$CONH), 2.15 (2H, m, C*H*$_2$CONH), 3.36-3.40 (2H, m, NHC*H*$_2$), 3.49-3.52 (2H, m, NHCH$_2$C*H*$_2$), 3.51-3.58 (4H, m, OC*H*$_2$C*H*$_2$O), 3.79-3.83 (2H, m, C*H*$_2$CH$_2$N(1')), 3.90 and 3.97 (2H, s, C*H*$_2$COO), 4.13 (2H, q, *J* 7.2, OC*H*$_2$CH$_3$), 4.46-4.48 (2H, m, C*H*$_2$N(1')), 4.50 and 4.52 (2H, s, C(4')C*H*$_2$), 6.01 and 6.28 (1H, br t, N*H*), 7.61 and 7.78 (1H, s, C(5')H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.1 (OCH$_2$*C*H$_3$ and *C*H$_3$), 22.6 (*C*H$_2$), 25.7 (*C*H$_2$), 29.2 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 31.8 (*C*H$_2$), 36.6 (*C*H$_2$), 39.2 (NH*C*H$_2$), 42.8 and 43.2 (C(4')*C*H$_2$), 48.0 and 48.9 (*C*H$_2$COO), 50.0 (*C*H$_2$N(1')), 60.9 (O*C*H$_2$CH$_3$), 69.2 (O*C*H$_2$), 69.3 (O*C*H$_2$), 69.9 (O*C*H$_2$), 70.1 (O*C*H$_2$), 123.1 and 123.5 (C(5')H), 144.2 and 144.6 (C(4')), 169.8 (*C*OO), 173.2 (*C*ONH). Seven signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of ethyl 2-(thymin-1-ylacetyl){[1-(8-(pentacosa-10,12-diynamido)-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl]methyl}aminoacetic acid (125a)

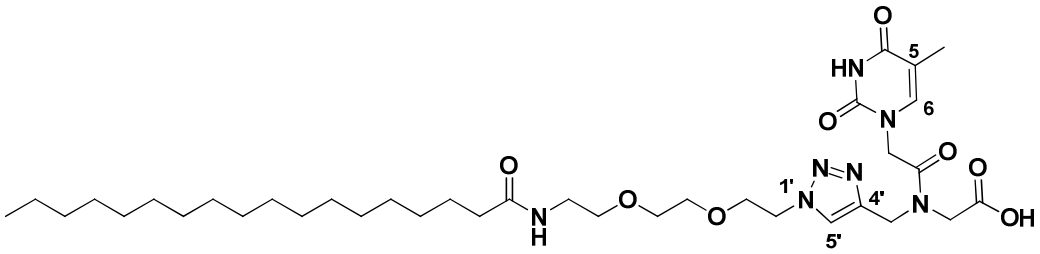


$C_{43}H_{67}N_7O_8$

95a (0.062 g, 0.074 mmol) was dissolved in tetrahydrofuran (5 mL) and 1 M lithium hydroxide solution (3 mL) was added. After stirring for 2 h, water (5 mL) was added and the aqueous layer separated and washed with dichloromethane (2 × 10 mL) before being acidified to pH 2 with 1 M hydrochloric acid solution. This acidic aqueous solution was then extracted with ethyl acetate (5 × 25 mL) and the combined organic phase was dried over magnesium sulphate anhydrous. After filtration, the solvent was evaporated under reduced pressure giving a crude solid which was purified by flash column chromatography (dichloromethane-methanol-acetic acid 89:10:1). The title compound (**125a**; 0.051 g, 0.063 mmol; 82%) was obtained as a sticky white solid; R_f (dichloromethane-methanol-acetic acid 89:10:1) = 0.16; υ_{max}(film)/cm^{-1} 3328br s, 2927br s, 2853s, 1767s, 1721s, 1679br s, 1545m; δ_H (400 MHz; DMSO-d_6) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.9, C*H*$_3$), 1.24-1.29 (32H, br s, 16 × C*H*$_2$), 1.71 (3H, s, C(5)-C*H*$_3$), 2.01 (2H, t, *J* 7.6, C*H*$_2$CONH), 2.22 (4H, t, *J* 7.1, 2 × C*H*$_2$C≡C), 3.14-3.16 (2H, m, NHC*H*$_2$), 3.30 (2H, m, NHCH$_2$C*H*$_2$), 3.43-3.51 (4H, m, OC*H*$_2$C*H*$_2$O), 3.73 (2H, m, C*H*$_2$CH$_2$N(1')), 4.13 and 4.27 (2H, s, C*H*$_2$COO), 4.42-4.49 (4H, m, C*H*$_2$CON and C*H*$_2$N(1')), 4.61 and 4.78 (2H, s, C(4')C*H*$_2$), 7.24 and 7.31 (1H, s, C(6)H), 7.91 and 8.12 (1H, s, C(5')H); δ_C (100 MHz; DMSO-d_6) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 11.9 (C(5)-*C*H$_3$), 13.9 (*C*H$_3$), 18.2 (*C*H$_2$C≡C), 22.7 (*C*H$_2$), 25.3 (*C*H$_2$CH$_2$CONH), 27.6 (*C*H$_2$), 27.7 (*C*H$_2$), 28.1 (*C*H$_2$), 28.2 (*C*H$_2$), 28.4 (*C*H$_2$), 28.6 (*C*H$_2$), 28.7 (*C*H$_2$), 28.8 (*C*H$_2$), 28.9 (*C*H$_2$), 29.0 (*C*H$_2$), 31.3 (*C*H$_2$), 35.6 (*C*H$_2$CONH), 38.7 (NH*C*H$_2$), 42.1 (C(4')*C*H$_2$), 47.8 (*C*H$_2$COO), 49.3 and 49.5 (*C*H$_2$CON), 50.2 and 50.4 (*C*H$_2$N(1')), 65.3 (*C*≡C-C≡*C*), 68.7 (*C*H$_2$CH$_2$N(1')), 69.2 (O*C*H$_2$), 69.4 (O*C*H$_2$), 69.5 (O*C*H$_2$), 77.9 (C≡*C*-*C*≡C), 108.0 (C(5)), 123.8 (C(5')H), 142.5 (C(6)H), 143.2

(C(4')), 151.1 (C(4)O), 164.4 (C(2)O), 167.3 (CH_2CON), 172.3 (*C*ONH), 172.3 (*C*OOH); *m/z* (ESI) 810 ($[M+H]^+$, 100%); HRMS (ESI) (Found: $[M+H]^+$, 810.5113. $C_{43}H_{68}N_7O_8$ requires *m/z*, 810.5124). Four signals corresponding to CH_2 from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of 2-(thymin-1-ylacetyl)((1-(8-stearamido-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl)methyl)aminoacetic acid (126a)

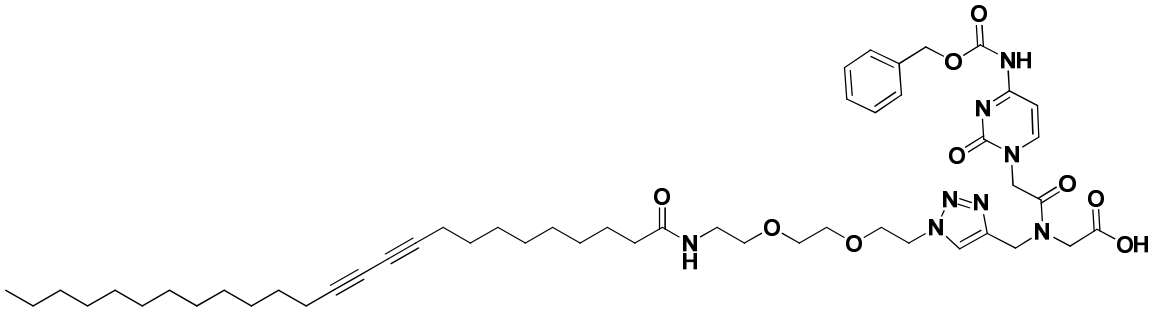


$C_{36}H_{61}N_7O_8$

96a (0.119 g, 0.160 mmol) was dissolved in tetrahydrofuran (5 mL) and 1 M lithium hydroxide solution (3 mL) was added. After stirring for 2 h, water (5 mL) was added and the aqueous layer separated and washed with dichloromethane (2 × 10 mL) before being acidified to pH 2 with 1 M hydrochloric acid solution. This acidic aqueous solution was then extracted with ethyl acetate (5 × 25 mL) and the combined organic phase was dried over magnesium sulphate anhydrous. After filtration, the solvent was evaporated under reduced pressure giving a crude solid which was purified by flash column chromatography (dichloromethane-methanol-acetic acid 89:10:1). The title compound (**126a**; 0.094 g, 0.13 mmol; 82%) was obtained as a white solid: m.p. 106-108 °C; R_f (dichloromethane-methanol-acetic acid 89:10:1) = 0.28; υ_{max}(film)/cm^{-1} 3217br m, 2926s, 2854s, 1711s, 1669br s, 1555w; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.3, C*H*$_3$), 1.12-1.33 (28H, m, 14 × C*H*$_2$), 1.58 (2H, m, C*H*$_2$CH$_2$CONH), 1.88 and 1.91 (3H, s, C(5)-C*H*$_3$), 2.15 (2H, t, *J* 7.6, C*H*$_2$CONH), 3.38-3.40 (2H, m, NHC*H*$_2$), 3.49-3.51 (2H, m, NHCH$_2$C*H*$_2$), 3.52-3.60 (4H, m, OC*H*$_2$C*H*$_2$O), 3.78-3.81 (2H, m, C*H*$_2$CH$_2$N(1')), 4.09 and 4.15 (2H, s, C*H*$_2$COO), 4.48 and 4.52 (2H, br t, C*H*$_2$N(1')), 4.49 and 4.82 (2H, s, C*H*$_2$CON), 4.69 and 4.71 (2H, s, C(4')C*H*$_2$), 6.45 (1H, br t, N*H*), 7.04 (1H, br s, C(6)H), 7.72 and 7.95 (1H, s, C(5')H), 9.66 and 9.72 (1H, br s, N(3)H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide

bond) 12.2 (C(5)-*C*H_3), 14.1 (*C*H_3), 22.7 (*C*H_2), 25.7 (*C*H_2CH_2CONH), 29.1 (*C*H_2), 29.3 (*C*H_2), 29.4 (*C*H_2), 29.5 (*C*H_2), 29.6 (*C*H_2), 29.7 (*C*H_2), 31.9 (*C*H_2), 36.6 (*C*H_2CONH), 39.1 (NH*C*H_2), 42.6 and 43.6 (C(4')*C*H_2), 47.6 and 49.1 (*C*H_2COO), 47.9 and 48.0 (*C*H_2CON), 50.1 and 50.2 (*C*H_2N(1')), 69.3 (*C*H_2CH_2N(1')), 69.8 (O*C*H_2), 69.9 (O*C*H_2), 70.3 (O*C*H_2), 110.8 and 110.9 (C(5)), 124.5 and 124.6 (C(5')H), 141.2 (C(6)H), 142.2 and 142.4 (C(4')), 151.2 (C(4)O), 164.5 (C(2)O), 167.1 and 167.3 (CH$_2$*C*ON), 174.2 (*C*ONH), 174.3 (*C*OOH); *m/z* (ESI) 720 ([M+H]$^+$, 100%); HRMS (ESI) (Found: [M+H]$^+$, 720.4639. $C_{36}H_{62}N_7O_8$ requires *m/z*, 720.4654). Six signals corresponding to *C*H_2 from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

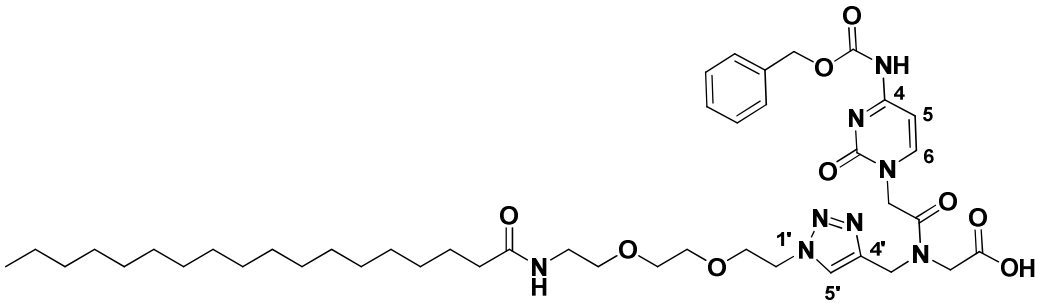
Attempted preparation of ethyl 2-(*N*^{4}-(benzyloxycarbonyl)-cytosin-1-ylacetyl) {[1-(8-(pentacosa-10,12-diynamido)-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methy} aminoacetic acid (125b)



$C_{50}H_{72}N_8O_9$

Compound **95b** (0.25 g, 0.27 mmol) was dissolved in tetrahydrofuran (20 mL) and 1 M lithium hydroxide solution (10 mL) was added. The milky white solution was stirred at room temperature for 1 h. After this time water (40 mL) was added and the mixture was washed with dichloromethane (2 × 40 mL). The aqueous phase was then acidified to pH 3 using 0.7 M citric acid solution and extracted with dichloromethane (9 × 80 mL). The combined organic phase was dried over anhydrous magnesium sulphate and filtrated. The crude residue was coevaporated with toluene (3 × 10 mL) and diethyl ether (3 × 10 mL). The analysis of the sticky white solid obtained did not match the title compound expected.

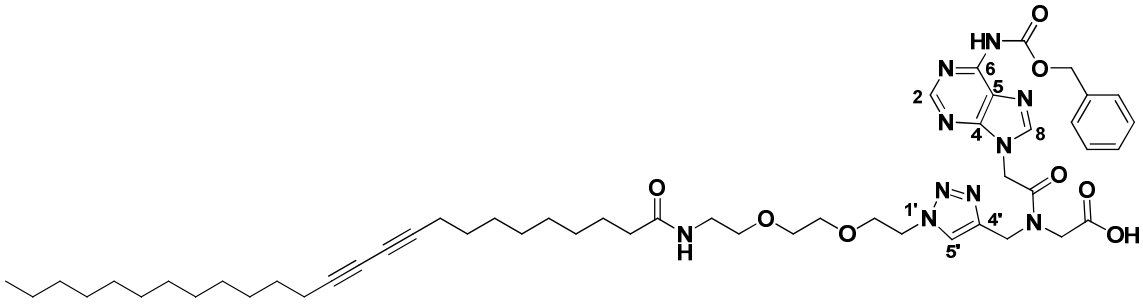
Attempted preparation of ethyl 2-[*N*^{4}-(benzyloxycarbonyl)cytosin-1-ylacetyl]{[1-(8-stearamido-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl]methyl}aminoacetic acid (126b)



$C_{43}H_{66}N_8O_9$

96b (0.093 g, 0.11 mmol) was dissolved in tetrahydrofuran (10 mL) and 1 M lithium hydroxide solution (5 mL) was added. The milky white solution was stirred at room temperature for 3 h. After this time water (20 mL) was added and the mixture was washed with dichloromethane (2 × 20 mL). The aqueous phase was then acidified to pH 1 using 0.7 M citric acid solution and extracted with ethyl acetate (8 × 25 mL). The combined organic phase was dried over anhydrous magnesium sulphate and filtrated. After evaporation of the solvent, a white solid was obtained but further analysis failed confirming the structure of the product.

Preparation of 2-[*N*^{6}-(benzyloxycarbonyl)-adenine-9-ylacetyl]{[1-(8-(pentacosa-10,12-diynamido)-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}aminoacetic acid (125c)

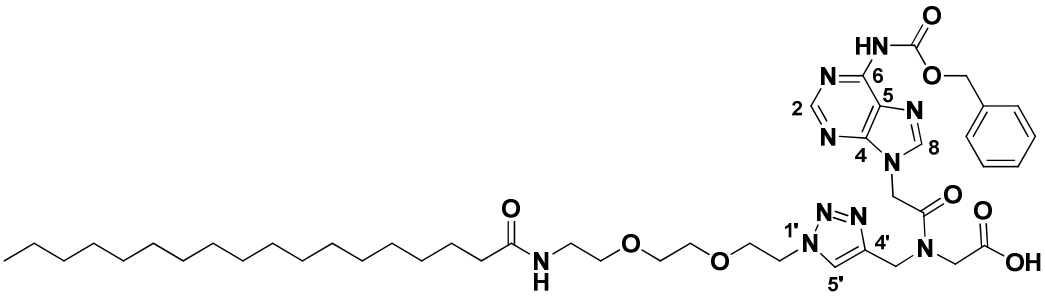


$C_{51}H_{72}N_{10}O_8$

95c (0.13 g, 0.13 mmol) was dissolved in dioxane (1.2 mL) and 0.77 M sodium hydroxide solution (0.42 mL) was added. After stirring for 30 min, water (2 mL) was added and the aqueous layer separated and washed with dichloromethane (3 × 5 mL) before

being acidified to pH 3 with 2 M citric acid solution. This acidic aqueous solution was then extracted with ethyl acetate (7 × 2 mL) and the combined organic phase was dried over magnesium sulphate anhydrous. After filtration, the solvent was evaporated under reduced pressure giving a crude solid which was purified by flash column chromatography (dichloromethane-methanol-acetic acid 89:10:1). The title compound (**125c**; 0.11 g, 0.12 mmol; 92%) was obtained as a sticky off-white solid; R_f (dichloromethane-methanol-acetic acid 89:10:1) = 0.21; υ_{max}(film)/cm^{-1} 3329br s, 2927br s, 2853s, 1772s, 1721s, 1679br s, 1545m; 1545m δ_H (200 MHz; $CDCl_3$ + 1 drop MeOH-*d*$_4$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.88 (3H, t, *J* 6.9, C*H*$_3$), 1.17-1.35 (26H, m, 13 × C*H*$_2$), 1.40-1.48 (4H, m, 2 × C*H*$_2$CH$_2$C≡C), 1.50-1.57 (2H, m, C*H*$_2$CH$_2$CONH), 2.05-2.12 (2H, m, C*H*$_2$CONH), 2.15-2.20 (4H, m, 2 × C*H*$_2$C≡C), 3.27-3.35 (2H, m, NHC*H*$_2$), 3.39-3.58 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.79 and 3.83 (2H, t, *J* 4.8, C*H*$_2$CH$_2$N(1')), 4.09 and 4.31 (2H, s, C*H*$_2$COO), 4.43 and 4.52 (2H, t, *J* 4.8, C*H*$_2$N(1')), 4.63 and 4.76 (2H, s, C*H*$_2$CON), 5.03 and 5.40 (2H, s, C(4')C*H*$_2$), 5.23 (2H, s, OC*H*$_2$Ar), 7.27-7.41 (5H, m, 5 × Ar-CH), 7.56 and 7.98 (1H, s, C(5')H), 8.05 and 8.08 (1H, s, C(8)H), 8.61 and 8.65 (1H, s, C(2)H); δ_C (400 MHz; $CDCl_3$ + 1 drop MeOH-*d*$_4$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 13.1 (*C*H$_3$), 19.2 (*C*H$_2$C≡C), 22.6 (*C*H$_2$), 28.3 (*C*H$_2$CH$_2$CO), 28.2 (*C*H$_2$CH$_2$C≡C), 28.7 (*C*H$_2$), 28.9 (*C*H$_2$), 29.2 (*C*H$_2$), 29.3 (*C*H$_2$), 29.4 (*C*H$_2$), 29.5 (*C*H$_2$), 29.6 (*C*H$_2$), 29.8 (*C*H$_2$), 31.8 (*C*H$_2$), 36.3 (*C*H$_2$CONH), 38.9 (NH*C*H$_2$), 42.6 and 43.5 (*C*H$_2$CON), 43.6 and 44.1 (C(4')*C*H$_2$), 47.8 and 48.8 (*C*H$_2$COO), 50.2 and 50.4 (*C*H$_2$N(1')), 65.1 and 65.2 (*C*≡C-C≡*C*), 67.5 (O*C*H$_2$Ar), 69.7 and 69.8 (*C*H$_2$CH$_2$N(1')), 69.9 (O*C*H$_2$), 70.0 (O*C*H$_2$), 70.4 (O*C*H$_2$), 77.2 and 77.5 (C≡*C*-*C*≡C), 121.2 (C(5)), 124.3 and 124.5 (C(5')H), 128.5 (Ar-CH), 128.6 (Ar-CH), 128.7 (Ar-CH), 135.4 (Ar-C), 142.1 and 142.3 (C(4')), 143.8 and 144.0 (C(8)H), 149.5 (C(4)), 151.2 (NH*C*OO), 151.5 (C(6)), 152.6 (C(2)H), 166.5 and 166.7 (CH$_2$*C*ON), 170.1 (*C*ONH), 174.1 (*C*OOH); *m/z* (ESI) 953 ([M+H]$^+$, 100%), 330 (7); HRMS (ESI) (Found: [M+H]$^+$, 953.5603. $C_{51}H_{73}N_{10}O_8$ requires *m/z*, 953.5607). Five signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of 2-[N^6-(benzyloxycarbonyl)adenin-9-ylacetyl]{[1-(8-stearamido-3,6-dioxaoctanyl)-1*H*-1,2,3-triazol-4-yl]methyl}aminoacetic acid (126c)

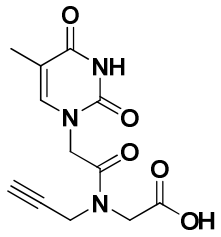


$C_{44}H_{66}N_{10}O_8$

Compound **96c** (0.10 g, 0.11 mmol) was dissolved in dioxane (1 mL) and 0.77 M sodium hydroxide solution (0.5 mL) was added. After stirring for 30 min, water (2 mL) was added and the aqueous layer separated and washed with dichloromethane (2 × 5 mL) before being acidified to pH 3 with 2 M citric acid solution. This acidic aqueous solution was then extracted with ethyl acetate (7 × 5 mL) and the combined organic phase was dried over anhydrous magnesium sulphate. After filtration, the solvent was evaporated under reduced pressure giving a crude solid which was purified by flash column chromatography (dichloromethane-methanol-acetic acid 89:10:1). The title compound (**126c**; 0.050 g, 0.058 mmol; 53%) was obtained as a white fluffy solid: R_f (dichloromethane-methanol-acetic acid 89:10:1) = 0.28; υ_{max}(film)/cm^{-1} 3217w, 2922s, 2852s, 1745m, 1665m, 1613m; δ_H (400 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 0.83 (3H, t, *J* 6.3, C*H*$_3$), 1.15-1.30 (28H, br s, 14 × C*H*$_2$), 1.48-1.59 (2H, m, C*H*$_2$CH$_2$CONH), 2.09-2.17 (2H, m, C*H*$_2$CONH), 3.25-3.32 (2H, m, NHC*H*$_2$), 3.40-3.59 (6H, m, C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.81 and 3.88 (2H, t, *J* 4.8, C*H*$_2$CH$_2$N(1')), 4.10 and 4.23 (2H, s, C*H*$_2$COO), 4.48 and 4.57 (2H, t, *J* 4.8, C*H*$_2$N(1')), 4.65 and 4.82 (2H, s, C*H*$_2$CON), 5.13 and 5.50 (2H, s, C(4')C*H*$_2$), 5.26 (2H, s, OC*H*$_2$Ar), 7.22-7.43 (5H, m, 5 × Ar-CH), 7.79 and 8.11 (1H, s, C(5')H), 8.16 and 8.22 (1H, s, C(8)H), 8.57 and 8.60 (1H, s, C(2)H); δ_C (100 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 14.3 (*C*H$_3$), 22.9 (*C*H$_2$), 26.1 (*C*H$_2$CH$_2$CON), 29.6 (*C*H$_2$), 29.7 (*C*H$_2$), 29.8 (*C*H$_2$), 29.9 (*C*H$_2$), 30.0 (*C*H$_2$), 30.1 (*C*H$_2$), 32.2 (*C*H$_2$), 36.7 (*C*H$_2$CON), 39.4 (NH*C*H$_2$), 43.1 and 43.6 (*C*H$_2$CON), 44.6 and 44.9 (C(4')*C*H$_2$), 48.0 and 48.9 (*C*H$_2$COO), 50.5 and 50.6 (*C*H$_2$N(1')), 67.9 (O*C*H$_2$Ar), 69.3 and 69.4 (*C*H$_2$CH$_2$N(1')), 69.9 (O*C*H$_2$), 70.2 (O*C*H$_2$), 70.6 (O*C*H$_2$), 121.1 (C(5)), 124.8 and 125.0 (C(5')H), 128.5 (Ar-CH), 128.6

(Ar-CH), 128.7 (Ar-CH), 135.9 (Ar-C), 142.2 and 142.4 (C(4')), 144.9 and 145.0 (C(8)H), 149.7 (C(4)), 151.5 (NH*C*OOCH$_2$), 152.0 (C(6)), 152.6 (C(2)H), 167.1 and 167.3 (CH$_2$*C*ON), 171.5 (*C*ONH), 175.8 (*C*OOH); *m/z* (ESI) 863 ([M+H]$^+$, 100%), 835 (12), 382 (4), 330 (7), 278 (21); HRMS (ESI) (Found: [M+H]$^+$, 863.5138. $C_{46}H_{71}N_{10}O_8$ requires *m/z*, 863.5138). Six signals corresponding to *C*H$_2$ from the lipid chain are missing due to the overlap in the aliphatic region of ^{13}C NMR spectrum.

Preparation of *N*-(thymin-1-ylacetyl)-*N*-(prop-2-ynyl)glycine (127)

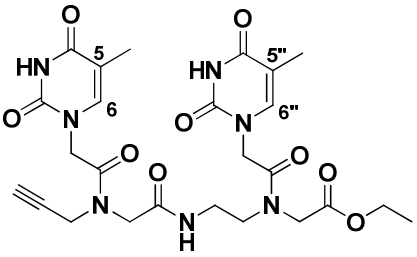


$C_{12}H_{13}N_3O_5$

84a (0.243 g, 0.790 mmol) was dissolved in dioxane (2.4 mL), to this was added 0.77 M sodium hydroxide solution (2.57 mL). The mixture was stirred at room temperature for 30 min. After this time, water (5.25 mL) was added and the solution was washed with dichloromethane (3 × 5 mL). The aqueous phase was then acidified to pH 3 using 2 M citric acid solution before being extracted with ethyl acetate (7 × 8 mL). The combined organic layers were washed with brine (8 mL), dried over anhydrous magnesium sulphate, filtered and evaporated to yield a crude residue that was subjected to flash column chromatography (dichloromethane-methanol-acetic acid 89:10:1). The title compound (**127**; 1.25 g, 3.25 mmol; 89%) was obtained as a sticky white solid; R_f (dichloromethane-methanol-acetic acid 89:10:1) = 0.21; δ_H (200 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 1.26 and 1.31 (2H, t, *J* 7.1, C*H*$_3$), 1.34 (3H, s, C(5)C*H*$_3$), 2.30 and 2.48 (1H, t, *J* 2.4, C*H*CCH$_2$), 4.18 and 4.26 (2H, q, *J* 7.1, OC*H*$_2$), 4.27-4.29 (4H, m, C*H*$_2$NC*H*$_2$), 4.50 and 4.70 (2H, s, C*H*$_2$CON), 7.01 (1H, s, C(6)H), 9.50 (1H, s, N(3)H); δ_C (50 MHz; $CDCl_3$) (2 rotational isomers observed due to restricted rotation about the tertiary amide bond) 12.2 (C(5)-*C*H$_3$), 14.0 and 14.1 (*C*H$_3$), 36.3 and 38.0 (HCC*C*H$_2$), 47.7 (*C*H$_2$COO), 48.0 and 48.1 (*C*H$_2$CON), 61.7 and 62.3 (O*C*H$_2$), 73.8 and 74.6 (HC*C*CH$_2$), 76.6 and 77.1 (*C*H), 110.9 (C(5)), 141.1 (C(6)H), 151.3 (C(2)O), 164.5 (C(4)O), 167.1 (CH$_2$*C*ON),

168.7 and 168.9 (*C*OO); *m/z* (ESI) 576 ([2M+NH_4]$^+$, 45%), 280 ([M+H]$^+$, 100); HRMS (ESI) (Found: [M+H]$^+$, 280.0931. $C_{12}H_{14}N_3O_5$ requires *m/z*, 280.0928).

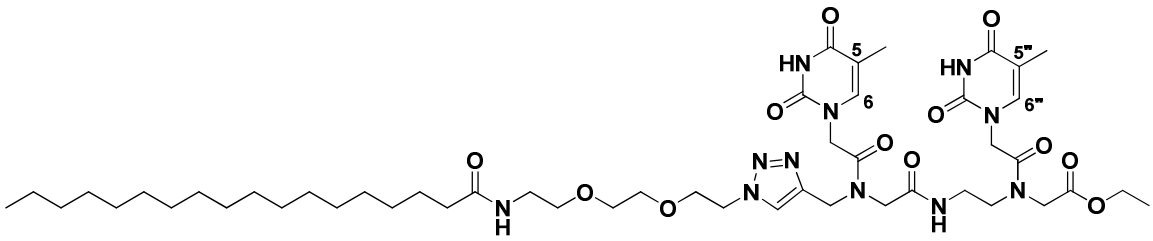
Preparation of *N*-{2-[(thymin-1-ylacetyl)-*N*-(prop-2-ynyl)amino]acetamido-ethyl}-*N*-(thymin-1-ylacetyl) glycine ethyl ester (128)



$C_{25}H_{31}N_7O_9$

127 (0.034 g, 0.12 mmol) in dimethylformamide (10 mL) was added HATU (0.050 g, 0.13 mmol) and DIPEA (0.025 mL, 0.14 mmol) and the solution was stirred for 15 min. After this time a solution of **82** (0.056 g, 0.13 mmol) in dimethylformamide (10 mL) was added dropwise, followed by DIPEA (0.100 mL, 0.58 mmol). The resulting mixture was stirred at room temperature for 20 h. After this time the solvents were removed *in vacuo*. The residue was dissolved in dichloromethane (10 mL) and the solution was washed successively with 1 M sodium bicarbonate (3 × 15 mL), 1 M potassium hydrogensulphate (2 × 15 mL), water (15 mL) and finally brine (20 mL). The organic layer was then dried over anhydrous magnesium sulphate and the solvent was evaporated under reduced pressure to yield a crude residue which was purified by flash column chromatography (ethyl acetate-methanol 90:10) to yield the title compound (**128**; 0.022 g, 0.04 mmol; 37%) as a white fluffy solid; R_f (ethyl acetate-methanol 90:10) = 0.15; δ_H (200 MHz; $CDCl_3$) 0.88 (3H, br t, C*H*$_3$), 1.85 (6H, s, C(5)-C*H*$_3$ and C(5")-C*H*$_3$), 2.37-2.42 (1H, m, *H*C≡C), 4.06-4.32 (12H, m, 6 × C*H*$_2$), 4.43 (2H, s, C*H*$_2$COO), 4.63 (2H, s, NC*H*$_2$CONH), 6.96-7.00 (2H, s, C(6)*H* and C(6")*H*), 9.25 (2H, br s, N(3)H); δ_C (50 MHz; $CDCl_3$) 11.8 (CH_2CH_3), 13.6 (C(5)-*C*H_3 and C(5")-*C*H_3), 36.1 (HC≡C*C*H_2), 37.1 (NH*C*H_2CH_2), 46.8-48.1 (5 × *C*H_2), 62.1 (O*C*H_2CH_3), 73.5 (HC≡*C*CH_2), 76.6 (H*C*≡CCH_2), 110.5 (C(5) and C(5")), 140.7 (C(6)*H* and C(6")*H*), 151.1 (C(4)O), 164.1 (C(2)O), 166.3 (CH_2*C*ON), 168.2 (*C*OO), 168.6 (*C*ONH); *m/z* (ESI) 591 ([M+NH_4]$^+$, 100), 573 ([M]$^+$, 62), 377 (13), 355 (29); HRMS (ESI) (Found: [M+ NH_4]$^+$, 591.2514. $C_{25}H_{35}N_8O_9$ requires *m/z*, 591.2522).

Preparation of *N*-{2-[(thymin-1-ylacetyl)-*N*-{[1-(8-stearamido-3,6-dioxaoctanyl)-*1H*-1,2,3-triazol-4-yl]methyl}amino]acetamido ethyl}-*N*-(thymin-1-ylacetyl) glycine ethyl ester (129)



$C_{49}H_{79}N_{11}O_{12}$

To the azide **99** (0.014 g, 0.032 mmol) in H_2O/t-BuOH/DCM (6 mL, 1:1:1) was added the dimer **128** (0.015 g, 0.026 mmol), followed by $CuSO_4.5H_2O$ (0.002 g, 30 mol %) and 1 M sodium ascorbate (0.016 mL, 62 mol %). The solution was stirred for 24 h at room temperature. After this time the aqueous phase was separated and the organic phase was concentrated. The crude residue was subjected to flash column chromatography (ethyl acetate - methanol 92:8) to give the title compound (**129**; 0.012 g, 0.012 mmol, 46%) as a white sticky solid; R_f (ethyl acetate-methanol 90:10) = 0.15; *m/z* (ESI) 1038 $[M+Na]^+$ (5%), 1015 ($[M+H]^+$, 100), 987 (13), 463 (2), 441 (13), 391 (3), 279 (2); HRMS (ESI) (Found: $[M+H]^+$, 1014.5977. $C_{49}H_{80}N_{11}O_{12}$ requires *m/z*, 1014.5982).

Preparation of liposome solution

The formation of liposomes was accomplished by first preparing solutions of 10,12-pentacosadiynoic acid (**5**) of known concentration in chloroform. The same step was carried out for the chosen lipid model. Then, the appropriate amounts of 10,12-pentacosadiynoic acid (**5**) and the chosen model were mixed in a test tube. Subsequently, the chloroform was evaporated using a gentle stream of nitrogen gas so that to obtain a thin (white) film on the glass. After addition of the appropriate amount of deionised water to give a total lipid concentration of 1 mM, the sample was heated to about 80 °C and sonicated for 15 min using a probe type sonicator (Sonozap®). The formed liposomes solutions obtained are opaque-transparent. The solution was then cooled overnight to 4 °C in a fridge. After warming to room temperature, the solution was purged with nitrogen gas for 5 min. The polymerisation was finally achieved by

placing the sample under a UV lamp (254 nm) for a variable amount of time (up to 30 min). The distance between the lamp and the sample was *ca.* 3 cm. The polymerized liposome solutions were stored for several months in glass vials at 4 °C in the fridge.

APPENDICES

APPENDIX A[§]

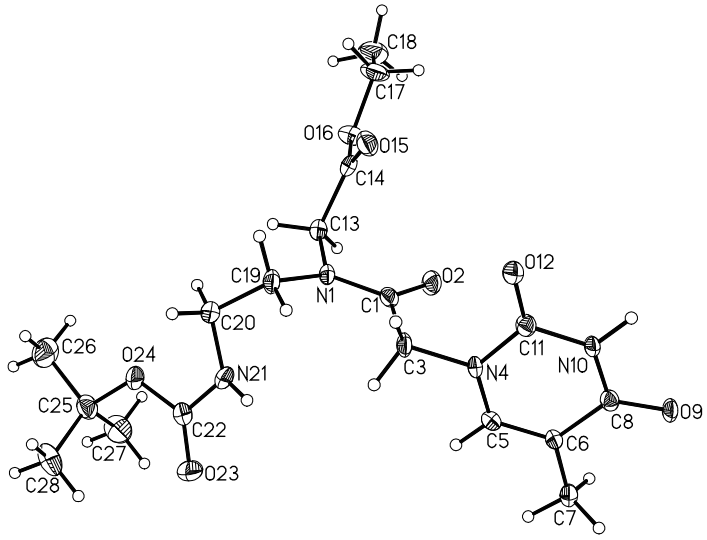


*Table 5: Crystal data and structure refinement for compound **70***

Empirical formula	$C_{18}H_{28}N_4O_7$	
Formula weight	412.44	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P2(1)/n	
Unit cell dimensions	a = 15.901(3) Å	α= 90°.
	b = 6.8936(15) Å	β= 100.450(12)°.
	c = 19.098(5) Å	γ = 90°.
Volume	2058.7(8) Å^3	
Z	4	
Density (calculated)	1.331 Mg/m^3	
Absorption coefficient	0.103 mm^{-1}	
F(000)	880	
Crystal size	0.60 x 0.24 x 0.08 mm^3	
Theta range for data collection	1.54 to 23.32°.	

[§] Data was collected by Dr Georgina Rosair on a Bruker X8 Apex2 CCD single crystal diffractometer at 100K.

Index ranges	-17<=h<=17, -7<=k<=7, -21<=l<=21
Reflections collected	26151
Independent reflections	2972 [R(int) = 0.0974]
Completeness to theta = 23.32°	99.5 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.992 and 0.574
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	2972 / 0 / 275
Goodness-of-fit on F^2	1.114
Final R indices [I>2sigma(I)]	R1 = 0.0487, wR2 = 0.1158
R indices (all data)	R1 = 0.1025, wR2 = 0.1596
Largest diff. peak and hole	0.291 and -0.422 e.$Å^{-3}$

*Table 6: Bond lengths [Å] and angles [°] for compound **70***

C(1)-O(2)	1.233(4)
C(1)-N(1)	1.360(5)
C(1)-C(3)	1.521(5)
N(1)-C(13)	1.441(5)
N(1)-C(19)	1.473(5)
C(3)-N(4)	1.461(4)
C(3)-H(3A)	0.9900
C(3)-H(3B)	0.9900
N(4)-C(11)	1.378(5)
N(4)-C(5)	1.382(5)
C(5)-C(6)	1.346(5)
C(5)-H(5)	0.9500
C(6)-C(8)	1.450(5)
C(6)-C(7)	1.492(5)
C(7)-H(7A)	0.9800
C(7)-H(7B)	0.9800
C(7)-H(7C)	0.9800
C(8)-O(9)	1.236(4)
C(8)-N(10)	1.383(5)
N(10)-C(11)	1.373(5)

N(10)-H(10)	0.98(4)
C(11)-O(12)	1.226(4)
C(13)-C(14)	1.512(5)
C(13)-H(13A)	0.9900
C(13)-H(13B)	0.9900
C(14)-O(15)	1.203(4)
C(14)-O(16)	1.343(5)
O(16)-C(17)	1.456(5)
C(17)-C(18)	1.492(6)
C(17)-H(17A)	0.9900
C(17)-H(17B)	0.9900
C(18)-H(18A)	0.9800
C(18)-H(18B)	0.9800
C(18)-H(18C)	0.9800
C(19)-C(20)	1.517(5)
C(19)-H(19A)	0.9900
C(19)-H(19B)	0.9900
C(20)-N(21)	1.454(5)
C(20)-H(20A)	0.9900
C(20)-H(20B)	0.9900
N(21)-C(22)	1.354(5)
N(21)-H(21)	0.79(4)
C(22)-O(23)	1.218(5)
C(22)-O(24)	1.350(5)
O(24)-C(25)	1.481(5)
C(25)-C(28)	1.514(6)
C(25)-C(27)	1.519(6)
C(25)-C(26)	1.520(6)
C(26)-H(26A)	0.9800
C(26)-H(26B)	0.9800
C(26)-H(26C)	0.9800
C(27)-H(27A)	0.9800
C(27)-H(27B)	0.9800
C(27)-H(27C)	0.9800
C(28)-H(28A)	0.9800

C(28)-H(28B)	0.9800
C(28)-H(28C)	0.9800

O(2)-C(1)-N(1)	121.4(3)
O(2)-C(1)-C(3)	120.5(3)
N(1)-C(1)-C(3)	118.1(3)
C(1)-N(1)-C(13)	116.3(3)
C(1)-N(1)-C(19)	125.2(3)
C(13)-N(1)-C(19)	117.0(3)
N(4)-C(3)-C(1)	110.1(3)
N(4)-C(3)-H(3A)	109.6
C(1)-C(3)-H(3A)	109.6
N(4)-C(3)-H(3B)	109.6
C(1)-C(3)-H(3B)	109.6
H(3A)-C(3)-H(3B)	108.1
C(11)-N(4)-C(5)	121.3(3)
C(11)-N(4)-C(3)	117.8(3)
C(5)-N(4)-C(3)	120.3(3)
C(6)-C(5)-N(4)	123.8(3)
C(6)-C(5)-H(5)	118.1
N(4)-C(5)-H(5)	118.1
C(5)-C(6)-C(8)	117.4(3)
C(5)-C(6)-C(7)	123.4(3)
C(8)-C(6)-C(7)	119.2(3)
C(6)-C(7)-H(7A)	109.5
C(6)-C(7)-H(7B)	109.5
H(7A)-C(7)-H(7B)	109.5
C(6)-C(7)-H(7C)	109.5
H(7A)-C(7)-H(7C)	109.5
H(7B)-C(7)-H(7C)	109.5
O(9)-C(8)-N(10)	119.6(3)
O(9)-C(8)-C(6)	124.7(3)
N(10)-C(8)-C(6)	115.8(3)
C(11)-N(10)-C(8)	126.9(3)
C(11)-N(10)-H(10)	117(2)

C(8)-N(10)-H(10)	116(2)
O(12)-C(11)-N(10)	122.7(3)
O(12)-C(11)-N(4)	122.5(3)
N(10)-C(11)-N(4)	114.7(3)
N(1)-C(13)-C(14)	111.9(3)
N(1)-C(13)-H(13A)	109.2
C(14)-C(13)-H(13A)	109.2
N(1)-C(13)-H(13B)	109.2
C(14)-C(13)-H(13B)	109.2
H(13A)-C(13)-H(13B)	107.9
O(15)-C(14)-O(16)	124.3(4)
O(15)-C(14)-C(13)	126.3(4)
O(16)-C(14)-C(13)	109.3(3)
C(14)-O(16)-C(17)	116.8(3)
O(16)-C(17)-C(18)	106.4(3)
O(16)-C(17)-H(17A)	110.5
C(18)-C(17)-H(17A)	110.5
O(16)-C(17)-H(17B)	110.5
C(18)-C(17)-H(17B)	110.5
H(17A)-C(17)-H(17B)	108.6
C(17)-C(18)-H(18A)	109.5
C(17)-C(18)-H(18B)	109.5
H(18A)-C(18)-H(18B)	109.5
C(17)-C(18)-H(18C)	109.5
H(18A)-C(18)-H(18C)	109.5
H(18B)-C(18)-H(18C)	109.5
N(1)-C(19)-C(20)	114.7(3)
N(1)-C(19)-H(19A)	108.6
C(20)-C(19)-H(19A)	108.6
N(1)-C(19)-H(19B)	108.6
C(20)-C(19)-H(19B)	108.6
H(19A)-C(19)-H(19B)	107.6
N(21)-C(20)-C(19)	110.7(3)
N(21)-C(20)-H(20A)	109.5
C(19)-C(20)-H(20A)	109.5

N(21)-C(20)-H(20B)	109.5
C(19)-C(20)-H(20B)	109.5
H(20A)-C(20)-H(20B)	108.1
C(22)-N(21)-C(20)	124.7(3)
C(22)-N(21)-H(21)	111(3)
C(20)-N(21)-H(21)	115(3)
O(23)-C(22)-O(24)	125.3(4)
O(23)-C(22)-N(21)	123.7(4)
O(24)-C(22)-N(21)	110.9(4)
C(22)-O(24)-C(25)	119.9(3)
O(24)-C(25)-C(28)	110.6(3)
O(24)-C(25)-C(27)	109.9(3)
C(28)-C(25)-C(27)	112.3(4)
O(24)-C(25)-C(26)	102.0(3)
C(28)-C(25)-C(26)	110.8(3)
C(27)-C(25)-C(26)	110.8(4)
C(25)-C(26)-H(26A)	109.5
C(25)-C(26)-H(26B)	109.5
H(26A)-C(26)-H(26B)	109.5
C(25)-C(26)-H(26C)	109.5
H(26A)-C(26)-H(26C)	109.5
H(26B)-C(26)-H(26C)	109.5
C(25)-C(27)-H(27A)	109.5
C(25)-C(27)-H(27B)	109.5
H(27A)-C(27)-H(27B)	109.5
C(25)-C(27)-H(27C)	109.5
H(27A)-C(27)-H(27C)	109.5
H(27B)-C(27)-H(27C)	109.5
C(25)-C(28)-H(28A)	109.5
C(25)-C(28)-H(28B)	109.5
H(28A)-C(28)-H(28B)	109.5
C(25)-C(28)-H(28C)	109.5
H(28A)-C(28)-H(28C)	109.5
H(28B)-C(28)-H(28C)	109.5

Symmetry transformations used to generate equivalent atoms:

*Table 7: Torsion angles [°] for compound **70***

O(2)-C(1)-N(1)-C(13)	-5.9(5)
C(3)-C(1)-N(1)-C(13)	176.1(3)
O(2)-C(1)-N(1)-C(19)	-171.7(3)
C(3)-C(1)-N(1)-C(19)	10.3(5)
O(2)-C(1)-C(3)-N(4)	-0.4(5)
N(1)-C(1)-C(3)-N(4)	177.6(3)
C(1)-C(3)-N(4)-C(11)	83.6(4)
C(1)-C(3)-N(4)-C(5)	-87.4(4)
C(11)-N(4)-C(5)-C(6)	2.7(5)
C(3)-N(4)-C(5)-C(6)	173.4(3)
N(4)-C(5)-C(6)-C(8)	-1.8(5)
N(4)-C(5)-C(6)-C(7)	178.7(3)
C(5)-C(6)-C(8)-O(9)	-177.5(3)
C(7)-C(6)-C(8)-O(9)	2.0(5)
C(5)-C(6)-C(8)-N(10)	1.9(5)
C(7)-C(6)-C(8)-N(10)	-178.6(3)
O(9)-C(8)-N(10)-C(11)	176.2(3)
C(6)-C(8)-N(10)-C(11)	-3.2(5)
C(8)-N(10)-C(11)-O(12)	-177.2(3)
C(8)-N(10)-C(11)-N(4)	3.9(5)
C(5)-N(4)-C(11)-O(12)	177.7(3)
C(3)-N(4)-C(11)-O(12)	6.8(5)
C(5)-N(4)-C(11)-N(10)	-3.4(5)
C(3)-N(4)-C(11)-N(10)	-174.3(3)
C(1)-N(1)-C(13)-C(14)	-70.1(4)
C(19)-N(1)-C(13)-C(14)	96.9(4)
N(1)-C(13)-C(14)-O(15)	-15.5(5)
N(1)-C(13)-C(14)-O(16)	165.8(3)
O(15)-C(14)-O(16)-C(17)	-2.9(5)
C(13)-C(14)-O(16)-C(17)	175.8(3)
C(14)-O(16)-C(17)-C(18)	179.9(3)
C(1)-N(1)-C(19)-C(20)	-126.7(4)
C(13)-N(1)-C(19)-C(20)	67.6(4)

N(1)-C(19)-C(20)-N(21)	62.3(4)
C(19)-C(20)-N(21)-C(22)	-176.4(4)
C(20)-N(21)-C(22)-O(23)	-165.8(4)
C(20)-N(21)-C(22)-O(24)	15.8(6)
O(23)-C(22)-O(24)-C(25)	1.5(6)
N(21)-C(22)-O(24)-C(25)	179.8(3)
C(22)-O(24)-C(25)-C(28)	60.3(4)
C(22)-O(24)-C(25)-C(27)	-64.3(4)
C(22)-O(24)-C(25)-C(26)	178.2(3)

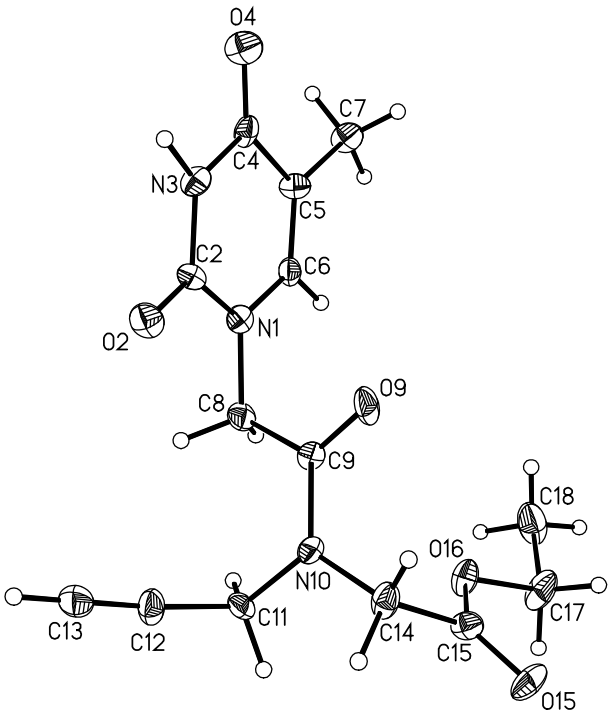
Symmetry transformations used to generate equivalent atoms:

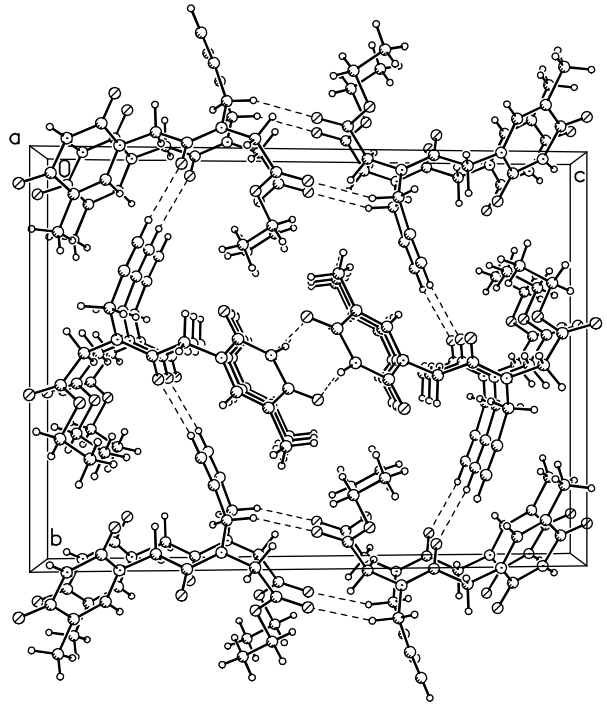
*Table 8: Hydrogen bonds for compound **70** [Å and °].*

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(10)-H(10)...O(9)#1	0.98(4)	1.86(4)	2.838(4)	174(3)
N(21)-H(21)...O(12)#2	0.79(4)	2.35(4)	3.079(5)	153(4)
N(21)-H(21)...O(15)#3	0.79(4)	2.83(4)	3.349(4)	125(3)

Symmetry transformations used to generate equivalent atoms:
#1 -x+2,-y,-z+1 #2 -x+3/2,y+1/2,-z+1/2 #3 x,y+1,z

APPENDIX B[**]





[**] Data was collected by Dr Georgina Rosair on a Bruker X8 Apex2 CCD single crystal diffractometer at 100K.

*Table 9 : Crystal data and structure refinement for compound **84a***

Empirical formula	$C_{14} H_{17} N_3 O_5$	
Formula weight	307.31	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P2(1)/n	
Unit cell dimensions	a = 4.8969(15) Å	α= 90°
	b = 15.069(5) Å	β= 96.444(13)°
	c = 20.118(5) Å	γ = 90°
Volume	1475.2(8) $Å^3$	
Z	4	
Density (calculated)	1.384 Mg/m^3	
Absorption coefficient	0.107 mm^{-1}	
F(000)	648	
Crystal size	0.44 x 0.20 x 0.12 mm^3	
Theta range for data collection	1.69 to 33.12°.	
Index ranges	-6<=h<=6, 0<=k<=21, 0<=l<=28	
Reflections collected	31957	
Independent reflections	6683 [R(int) = 0.0763]	
Completeness to theta = 25.00°	89.5 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.9873 and 0.6355	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	6683 / 19 / 206	
Goodness-of-fit on F^2	0.996	
Final R indices [I>2sigma(I)]	R1 = 0.0922, wR2 = 0.2518	
R indices (all data)	R1 = 0.1223, wR2 = 0.2700	
Largest diff. peak and hole	0.395 and -0.518 $e.Å^{-3}$	

*Table 10 : Bond lengths [Å] and angles [°] for compound **84a***

N(1)-C(2)	1.373(6)
N(1)-C(6)	1.394(6)
N(1)-C(8)	1.447(6)
C(2)-O(2)	1.222(5)
C(2)-N(3)	1.382(5)
N(3)-C(4)	1.377(6)
N(3)-H(3N)	0.899(19)
C(4)-O(4)	1.237(5)
C(4)-C(5)	1.454(7)
C(5)-C(6)	1.339(6)
C(5)-C(7)	1.508(6)
C(6)-H(6)	0.9500
C(7)-H(7A)	0.9800
C(7)-H(7B)	0.9800
C(7)-H(7C)	0.9800
C(8)-C(9)	1.518(7)
C(8)-H(8A)	0.9900
C(8)-H(8B)	0.9900
C(9)-O(9)	1.230(6)
C(9)-N(10)	1.368(6)
N(10)-C(14)	1.455(6)
N(10)-C(11)	1.460(6)
C(11)-C(12)	1.475(7)
C(11)-H(11A)	0.9900
C(11)-H(11B)	0.9900
C(12)-C(13)	1.191(7)
C(13)-H(13)	0.9500
C(14)-C(15)	1.517(7)
C(14)-H(14A)	0.9900
C(14)-H(14B)	0.9900
C(15)-O(15)	1.201(5)
C(15)-O(16)	1.321(5)
O(16)-C(17)	1.490(6)
C(17)-C(18)	1.472(7)

C(17)-H(17A)	0.9900
C(17)-H(17B)	0.9900
C(18)-H(18A)	0.9800
C(18)-H(18B)	0.9800
C(18)-H(18C)	0.9800

C(2)-N(1)-C(6)	121.2(4)
C(2)-N(1)-C(8)	117.8(4)
C(6)-N(1)-C(8)	120.8(4)
O(2)-C(2)-N(1)	122.5(4)
O(2)-C(2)-N(3)	122.5(4)
N(1)-C(2)-N(3)	115.0(4)
C(4)-N(3)-C(2)	127.0(4)
C(4)-N(3)-H(3N)	115(3)
C(2)-N(3)-H(3N)	118(3)
O(4)-C(4)-N(3)	120.8(4)
O(4)-C(4)-C(5)	123.9(4)
N(3)-C(4)-C(5)	115.3(4)
C(6)-C(5)-C(4)	118.4(4)
C(6)-C(5)-C(7)	123.0(4)
C(4)-C(5)-C(7)	118.5(4)
C(5)-C(6)-N(1)	123.1(4)
C(5)-C(6)-H(6)	118.5
N(1)-C(6)-H(6)	118.5
C(5)-C(7)-H(7A)	109.5
C(5)-C(7)-H(7B)	109.5
H(7A)-C(7)-H(7B)	109.5
C(5)-C(7)-H(7C)	109.5
H(7A)-C(7)-H(7C)	109.5
H(7B)-C(7)-H(7C)	109.5
N(1)-C(8)-C(9)	110.1(4)
N(1)-C(8)-H(8A)	109.6
C(9)-C(8)-H(8A)	109.6
N(1)-C(8)-H(8B)	109.6
C(9)-C(8)-H(8B)	109.6

H(8A)-C(8)-H(8B)	108.1
O(9)-C(9)-N(10)	122.6(4)
O(9)-C(9)-C(8)	121.4(4)
N(10)-C(9)-C(8)	116.0(4)
C(9)-N(10)-C(14)	119.1(4)
C(9)-N(10)-C(11)	122.6(4)
C(14)-N(10)-C(11)	118.3(4)
N(10)-C(11)-C(12)	111.8(4)
N(10)-C(11)-H(11A)	109.3
C(12)-C(11)-H(11A)	109.3
N(10)-C(11)-H(11B)	109.3
C(12)-C(11)-H(11B)	109.3
H(11A)-C(11)-H(11B)	107.9
C(13)-C(12)-C(11)	175.1(5)
C(12)-C(13)-H(13)	180.0
N(10)-C(14)-C(15)	114.6(4)
N(10)-C(14)-H(14A)	108.6
C(15)-C(14)-H(14A)	108.6
N(10)-C(14)-H(14B)	108.6
C(15)-C(14)-H(14B)	108.6
H(14A)-C(14)-H(14B)	107.6
O(15)-C(15)-O(16)	125.3(4)
O(15)-C(15)-C(14)	122.9(4)
O(16)-C(15)-C(14)	111.7(4)
C(15)-O(16)-C(17)	116.3(3)
C(18)-C(17)-O(16)	106.8(4)
C(18)-C(17)-H(17A)	110.4
O(16)-C(17)-H(17A)	110.4
C(18)-C(17)-H(17B)	110.4
O(16)-C(17)-H(17B)	110.4
H(17A)-C(17)-H(17B)	108.6
C(17)-C(18)-H(18A)	109.5
C(17)-C(18)-H(18B)	109.5
H(18A)-C(18)-H(18B)	109.5
C(17)-C(18)-H(18C)	109.5

H(18A)-C(18)-H(18C)	109.5
H(18B)-C(18)-H(18C)	109.5

*Table 11: Torsion angles [°] for compound **96a***

C(6)-N(1)-C(2)-O(2)	-174.9(4)
C(8)-N(1)-C(2)-O(2)	-0.9(6)
C(6)-N(1)-C(2)-N(3)	3.7(6)
C(8)-N(1)-C(2)-N(3)	177.7(4)
O(2)-C(2)-N(3)-C(4)	176.2(4)
N(1)-C(2)-N(3)-C(4)	-2.4(6)
C(2)-N(3)-C(4)-O(4)	178.7(4)
C(2)-N(3)-C(4)-C(5)	-0.3(7)
O(4)-C(4)-C(5)-C(6)	-177.2(5)
N(3)-C(4)-C(5)-C(6)	1.8(6)
O(4)-C(4)-C(5)-C(7)	2.7(7)
N(3)-C(4)-C(5)-C(7)	-178.3(4)
C(4)-C(5)-C(6)-N(1)	-0.5(7)
C(7)-C(5)-C(6)-N(1)	179.6(4)
C(2)-N(1)-C(6)-C(5)	-2.4(7)
C(8)-N(1)-C(6)-C(5)	-176.2(4)
C(2)-N(1)-C(8)-C(9)	-79.0(5)
C(6)-N(1)-C(8)-C(9)	95.0(5)
N(1)-C(8)-C(9)-O(9)	-9.2(6)
N(1)-C(8)-C(9)-N(10)	173.0(4)
O(9)-C(9)-N(10)-C(14)	-3.2(6)
C(8)-C(9)-N(10)-C(14)	174.6(4)
O(9)-C(9)-N(10)-C(11)	174.9(4)
C(8)-C(9)-N(10)-C(11)	-7.2(6)
C(9)-N(10)-C(11)-C(12)	-75.3(5)
C(14)-N(10)-C(11)-C(12)	102.8(4)
N(10)-C(11)-C(12)-C(13)	42(7)
C(9)-N(10)-C(14)-C(15)	-93.3(5)
C(11)-N(10)-C(14)-C(15)	88.5(5)
N(10)-C(14)-C(15)-O(15)	-164.9(4)
N(10)-C(14)-C(15)-O(16)	17.7(6)

O(15)-C(15)-O(16)-C(17)	-2.5(7)
C(14)-C(15)-O(16)-C(17)	174.9(4)
C(15)-O(16)-C(17)-C(18)	-167.5(4)

*Table 12: Hydrogen bonds for compound **84a** [Å and °].*

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(3)-H(3N)...O(4)#1	0.899(19)	1.90(2)	2.804(5)	179(5)

Symmetry transformations used to generate equivalent atoms:

#1 -x+1,-y+2,-z+2

REFERENCES

REFERENCES

1. K. L. Dueholm, M. Egholm, C. Behrens, L. Christensen, H. F. Hansen, T. Vulpius, K. H. Petersen, R. H. Berg, P. E. Nielsen and O. Buchardt, *J. Org. Chem.*, 1994, **59**, 5767-5773.
2. E. Murray, Ph.D. Thesis, Heriot-Watt University, 2004.
3. R. Dahm, *Dev. Biol.*, 2005, **278**, 274-288.
4. J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 737-738.
5. J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 964-967.
6. J. M. Berg, J. L. Tymoczko and L. Stryer, in *Biochemistry - 5th Ed.*, ed. W. H. F. a. Company, Michelle Julet, New York, Editon edn., 2002, pp. 747-750.
7. J. Wang, G. Rivas, X. Cai, M. Chicharro, C. Parrado, N. Dontha, A. Begleiter, M. Mowat, E. Palecek and P. E. Nielsen, *Anal. Chim. Acta*, 1997, **344**, 111-118.
8. J. Zhai, H. Cui and R. Yang, *Biotechnol. Adv.*, 1997, **15**, 43-58.
9. H. Cai, Y. Wang, P. He and Y. Fang, *Anal. Chim. Acta*, 2002, **469**, 165-172.
10. V. Pavlov, Y. Xiao, R. Gill, A. Dishon, M. Kotler and I. Willner, *Anal. Chem.*, 2004, **76**, 2152-2156.
11. S. Pan and L. Rothberg, *Langmuir*, 2005, **21**, 1022-1027.
12. G. Marrazza, I. Chianella and M. Mascini, *Biosens. Bioelectron.*, 1999, **14**, 43-51.
13. A. Sassolas, B. D. Leca-Bouvier and L. J. Blum, *Chem. Rev.* , 2008, **108**, 109-139.
14. A. B. Steel, T. M. Herne and M. J. Tarlov, *Anal. Chem.*, 1998, **70**, 4670-4677.
15. T. M. Herne and M. J. Tarlov, *J. Am. Chem. Soc.*, 1997, **119**, 8916-8920.
16. K. A. Peterlinz, R. M. Georgiadis, T. M. Herne and M. J. Tarlov, *J. Am. Chem. Soc.*, 1997, **119**, 3401-3402.
17. P. A. E. Piunno, U. J. Krull, R. H. E. Hudson, M. J. Damha and H. Cohen, *Anal. Chem.*, 1995, **67**, 2635-2643.
18. J. B. Lepecq and C. Paoletti, *J. Mol. Biol.*, 1967, **27**, 87-106.
19. H. Zipper, H. Brunner, J. Bernhagen and F. Vitzthum, *Nucleic Acids Res.*, 2004, **32**, 1-10.
20. R. M. Fryer, J. Randall, T. Yoshida, L. L. Hsiao, J. Blumenstock, K. E. Jensen, T. Dimofte, R. V. Jensen and S. R. Gullans, *Exp. Nephrol.*, 2002, **10**, 64-74.
21. W. Tan, K. Wang and T. J. Drake, *Curr. Op. Chem. Biol.*, 2004, **8**, 547-553.
22. N. E. Broude, *Trends Biotechnol.*, 2002, **20**, 249-256.
23. X. Liu and W. Tan, *Anal. Chem.*, 1999, **71**, 5054-5059.
24. X. Fang, X. Liu, S. Schuster and W. Tan, *J. Am. Chem. Soc.*, 1999, **121**, 2921-2922.
25. J. Spadavecchia, M. G. Manera, F. Quaranta, P. Siciliano and R. Rella, *Biosens. Bioelectron.*, 2005, **21**, 894-900.
26. H. J. Watts, D. Yeung and H. Parkes, *Anal. Chem.*, 1995, **67**, 4283-4289.
27. A. J. Thiel, A. G. Frutos, C. E. Jordan, R. M. Corn and L. M. Smith, *Anal. Chem.*, 1997, **69**, 4948-4956.
28. J. Wang, *Nucleic Acids Res.*, 2000, **28**, 3011-3016.
29. T. Liebermann and W. Knoll, *Colloids Surf., A*, 2000, **171**, 115-130.
30. T. Liebermann, W. Knoll, P. Sluka and R. Herrmann, *Colloids Surf., A*, 2000, **169**, 337-350.

31. J. Wang, *Chem. Eur. J.*, 1999, **5**, 1681-1685.
32. J. Wang, *Anal. Chim. Acta*, 2002, **469**, 63-71.
33. F. Azek, C. Grossiord, M. Joannes, B. Limoges and P. Brossier, *Anal. Biochem.*, 2000, **284**, 107-113.
34. S. Takenaka, Y. Uto, H. Saita, M. Yokoyama, H. Kondo and D. Wilson, *Chem. Commun.*, 1998, 1111-1112.
35. S. Takenaka, K. Yamashita, M. Takagi, Y. Uto and H. Kondo, *Anal. Chem.*, 2000, **72**, 1334-1341.
36. K. M. Millan, A. Saraullo and S. R. Mikkelsen, *Anal. Chem.*, 1994, **66**, 2943-2948.
37. J. Wang, G. Rivas, J. R. Fernandes, J. L. Lopez Paz, M. Jiang and R. Waymire, *Anal. Chim. Acta*, 1998, **375**, 197-203.
38. C. K. O'Sullivan and G. G. Guilbault, *Biosens. Bioelectron.*, 1999, **14**, 663-670.
39. K. M. Hansen, H.-F. Ji, G. Wu, R. Datar, R. Cote, A. Majumdar and T. Thundat, *Anal. Chem.*, 2001, **73**, 1567-1571.
40. P. E. Nielsen, M. Egholm and O. Buchardt, *Science*, 1991, **254**, 1497-1500.
41. P. E. Nielsen and G. Haaima, *Chem. Soc. Rev.* , 1997, **26**, 73-78.
42. M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895-1897.
43. M. Egholm, P. E. Nielsen, O. Buchardt and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 9677-9678.
44. M. Egholm, C. Behrens, L. Christensen, R. H. Berg, P. E. Nielsen and O. Buchardt, *J. Chem. Soc., Chem. Commun.*, 1993, 800-801.
45. M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566-568.
46. P. E. Nielsen and L. Christensen, *J. Am. Chem. Soc.*, 1996, **118**, 2287-2288.
47. P. E. Nielsen and M. Egholm, *Bioorg. Med. Chem.*, 2001, **9**, 2429-2434.
48. J. Lohse, O. Dahl and P. E. Nielsen, *P. Natl. Acad. Sci USA*, 1999, **96**, 11804-11808.
49. D. A. Braasch, C. J. Nulf and D. R. Corey, *Curr. Protoc. Nucleic Acid Chem.*, 2002, **Chapter 4**, Unit 4 11.
50. A. Ray and B. Norden, *FASEB J.*, 2000, **14**, 1041-1060.
51. K. Helle and P. E. Nielsen, *Nucleic Acids Res.*, 1996, **24**, 494-500.
52. L. Mologni, P. leCoutre, P. E. Nielsen and C. Gambacorti-Passerini, *Nucleic Acids Res.*, 1998, **26**, 1934-1938.
53. P. E. Nielsen, M. Egholm and O. Buchardt, *Gene*, 1994, **149**, 139-145.
54. R. Lee, N. Kaushik, M. J. Modak, R. Vinayak and V. N. Pandey, *Biochemistry*, 1998, **37**, 900-910.
55. U. Koppelhus and P. E. Nielsen, *Adv. Drug. Deliv. Rev.*, 2003, **55**, 267-280.
56. H. Orum, *Curr. Issues Mol. Bio.*, 2000, **2**, 27-30.
57. H. Orum, P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt and C. Stanley, *Nucleic Acids Res.*, 1993, **21**, 5332-5336.
58. S. Kwok, D. E. Kellogg, N. McKinney, D. Spasic, L. Goda, C. Levenson and J. J. Sninsky, *Nucleic Acids Res.*, 1990, **18**, 999-1005.

59. T. Iwamoto and S. Toshiaki, *Antimicrob. Agents Chemother.*, 2004, **48**, 4023-4026.
60. J. Wang, E. Palecek, P. E. Nielsen, G. Rivas, X. H. Cai, H. Shiraishi, N. Dontha, D. B. Luo and P. A. M. Farias, *J. Am. Chem. Soc.*, 1996, **118**, 7667-7670.
61. J. Wang, *Current Issues Molec. Biol.*, 1999, **1**, 117-122.
62. J. Wang, P. E. Nielsen, M. Jiang, X. Cai, J. R. Fernandes, D. H. Grant, M. Ozsoz, A. Beglieter and M. Mowat, *Anal. Chem.*, 1997, **69**, 5200-5202.
63. O. Brandt and J. D. Hoheisel, *Trends Biotechnol.* , 2004, **22**, 617-622.
64. G. Wegner, *Z. Naturforsch., B*, 1969, **24**, 824.
65. G. Wegner, *Macromol. Chem.*, 1972, **154**, 35-48.
66. G. Wegner, *Pure Appl. Chem.*, 1977, **49**, 443-4454.
67. R. H. Baughman, *J. Appl. Phys.*, 1972, **43**, 4362.
68. S. Okada, S. Peng, W. Spevak and D. Charych, *Acc. Chem. Res.*, 1998, **31**, 229-239.
69. M. A. Reppy and B. A. Pindzola, *Chem. Commun.*, 2007, **42**, 4317-4338.
70. B. Tieke and G. Lieser, *J. Colloid Interface Sci.*, 1982, **88**, 471-486.
71. H. Menzel, S. Horstmann, M. D. Mowery, M. Cai and C. E. Evans, *Polymer*, 2000, **41**, 8113-8119.
72. R. W. Carpick, D. Y. Sasaki, M. S. Marcus, M. A. Eriksson and A. R. Burns, *J. Phys.: Condens. Matter*, 2004, **16**, R679-R697.
73. F. Gaboriaud, R. Golan, R. Volinsky, A. Berman and R. Jelinek, *Langmuir*, 2001, **17**, 3651-3657.
74. R. Volinsky, F. Gaboriaud, A. Berman and R. Jelinek, *J. Phys. Chem. B*, 2002, **106**, 9231-9236.
75. R. R. C. New, in *Liposomes: a practical approach*, ed. R. R. C. New, Oxford University Press, Oxford, Editon edn., 1990, pp. 33-104.
76. S. Kolusheva, R. Kafri, M. Katz and R. Jelinek, *J. Am. Chem. Soc.*, 2001, **123**, 417-422.
77. S. Kolusheva, T. Shahal and R. Jelinek, *Biochemistry*, 2000, **39**, 15851-15859.
78. S. Kolusheva, E. Wachtel and R. Jelinek, *J. Lip. Res.*, 2003, **44**, 65-71.
79. R. Jelinek and S. Kolusheva, *Biotechnol. Adv.*, 2001, **19**, 109-118.
80. D. J. Ahn and J. M. Kim, *Acc. Chem. Res.*, 2008, **41**, 805-816.
81. R. R. Chance, *Macromolecules*, 1980, **13**, 396-398.
82. J. Yoon, S. K. Chae and J.-M. Kim, *J. Am. Chem. Soc.*, 2007, **129**, 3038-3039.
83. R. R. Chance, G. N. Patel and J. D. Witt, *J. Chem. Phys.*, 1979, **71**, 206-211.
84. D. J. Ahn, E.-H. Chae, G. S. Lee, H.-Y. Shim, T.-E. Chang, K.-D. Ahn and J.-M. Kim, *J. Am. Chem. Soc.*, 2003, **125**, 8976-8977.
85. A. Singh, R. B. Thompson and J. M. Schnur, *J. Am. Chem. Soc.*, 1986, **108**, 2785-2787.
86. Z. Yuan, C.-W. Lee and S.-H. Lee, *Angew. Chem. Int. Ed.*, 2004, **43**, 4197-4200.
87. H. W. Beckham and M. F. Rubner, *Macromolecules*, 1993, **26**, 5198-5201.
88. R. A. Nallicheri and M. F. Rubner, *Macromolecules*, 1991, **24**, 517-525.
89. S. Y. Okada, R. Jelinek and D. Charych, *Angew. Chem. Int. Ed.*, 1999, **38**, 655-659.
90. D. H. Charych, J. O. Nagy, W. Spevak and M. D. Bednarski, *Science*, 1993, **261**, 585-588.
91. A. Reichert, J. O. Nagy, W. Spevak and D. Charych, *J. Am. Chem. Soc.*, 1995, **117**, 829-830.
92. Z. Ma, J. Li, M. Liu, J. Cao, Z. Zou, J. Tu and L. Jiang, *J. Am. Chem. Soc.*, 1998, **120**, 12678-12679.

93. C. Wang and Z. Ma, *Anal. Bioanal.Chem.*, 2005, **382**, 1708-1710.
94. S. Kolusheva, T. Shahal and R. Jelinek, *J. Am. Chem. Soc.*, 2000, **122**, 776-780.
95. J.-S. Filhol, J. Deschamps, S. G. Dutremez, B. Boury, T. Barisien, L. Legrand and M. Schott, *J. Am. Chem. Soc.*, 2009, **131**, 6976-6988.
96. U. Jonas, K. Shah, S. Norvez and D. H. Charych, *J. Am. Chem. Soc.*, 1999, **121**, 4580-4588.
97. J.-M. Kim, Y. B. Lee, D. H. Yang, J.-S. Lee, G. S. Lee and D. J. Ahn, *J. Am. Chem. Soc.*, 2005, **127**, 17580-17581.
98. D. H. Charych, Q. Cheng, A. Reichert, G. Kuziemko, M. Stroh, J. O. Nagy, W. Spevak and R. C. Stevens, *Chem. Biol.*, 1996, **3**, 113-120.
99. Y. K. Jung, H. G. Park and J.-M. Kim, *Biosens. Bioelectron.*, 2006, **21**, 1536-1544.
100. S. Kolusheva, L. Boyer and R. Jelinek, *Nat. Biotechnol.*, 2000, **18**, 225-227.
101. N. Y. Lee, Y. K. Jung and H. G. Park, *Biochem. Eng. J.*, 2006, **29**, 103-108.
102. B. Ma, Y. Fan, L. Zhang, X. Kong, Y. Li and J. Li, *Colloids Surf., B*, 2003, **27**, 209-213.
103. J. J. Pan and D. Charych, *Langmuir*, 1997, **13**, 1365-1367.
104. R. Jelinek, S. Okada, S. Norvez and D. Charych, *Chem. Biol.*, 1998, **5**, 619-629.
105. R. Jelinek, *Drug Dev. Res.*, 2000, **50**, 497-501.
106. M. Katz, I. Ben-Shlush, S. Kolusheva and R. Jelinek, *Pharm. Res.*, 2006, **23**, 580-588.
107. Y. Amano and Q. Cheng, *Anal. Bioanal.Chem.*, 2005, **381**, 156-164.
108. S. Kashiwagi, *Nippon Rinsho*, 2003, **61**, 1963-1966.
109. G. M. Air and W. G. Laver, *Proteins: Struct. Funct. Bioinf.*, 1989, **6**, 341-356.
110. R. Maoz and J. Sagiv, *J. Colloid Interface Sci.*, 1984, **100**, 465-496.
111. J. Deng, Z. Sheng, K. Zhou, M. Duan, C. Y. Yu and L. Jiang, *Bioconjugate Chem.* , 2009, **20**, 533-537.
112. C. Wang, Z. Ma and Z. Su, *Sens. Actuators, B*, 2006, **113**, 510-515.
113. Y. K. Jung, T. W. Kim, J. Kim, J. M. Kim and H. G. Park, *Adv. Funct. Mater.*, 2008, **18**, 701-708.
114. A. Ghassemi and P. Rosenberg, *Biochem. Pharmacol.*, 1992, **44**, 1073-1083.
115. Y. Zhang, Y. Fan, C. Sun, D. Shen, Y. Li and J. Li, *Colloids Surf., B*, 2005, **40**, 137-142.
116. R. C. Stevens and Q. Cheng, *Langmuir*, 1998, **14**, 1974-1976.
117. C. Murray, Ph.D. Thesis, Heriot-Watt University, 1999.
118. W. E. Lindsell, C. Murray, P. N. Preston and T. A. J. Woodman, *Tetrahedron*, 2000, **56**, 1233-1245.
119. A. Farese, N. Patino, S. Dalleu and R. Guedj, *Tetrahedron Lett.*, 1996, **37**, 1413-1416.
120. L. Christensen, R. Fitzpatrick, B. Gildea, K. H. Petersen, H. F. Hansen, T. Koch, M. Egholm, O. Buchardt, P. E. Nielsen, J. Coull and et al., *J. Pept. Sci.*, 1995, **1**, 175-183.
121. J. O. Nagy, W. Spevak, D. H. Charych, M. E. Schaefer, J. H. Gilbert and M. D. Bednarski, *J. Cell. Biochem.*, 1993, 382-382.
122. W. Spevak, J. O. Nagy, D. H. Charych, M. E. Schaefer, J. H. Gilbert and M. D. Bednarski, *J. Am. Chem. Soc.*, 1993, **115**, 1146-1147.
123. J. P. Vernille, L. C. Kovell and J. W. Schneider, *Bioconjugate Chem.* , 2004, **15**, 1314-1321.
124. B. F. Marques and J. W. Schneider, *Langmuir*, 2005, **21**, 2488-2494.

125. J. W. Schneider, C. Lau, B. F. Marques and S. T. Grosser, *Abstr. Am. Chem. Soc.*, 2005, **229**, U716-U716.

126. A. M. P. Koskinen, T. Valo, S. Vihavainen and J. M. L. Hakala, *Bioorg. Med. Chem. Lett.*, 1995, **5**, 573-578.

127. F. G. Bordwell and H. E. Fried, *J. Org. Chem.*, 1991, **56**, 4218-4223.

128. E. E. Dueno, F. Chu, S.-I. Kim and K. W. Jung, *Tetrahedron Lett.*, 1999, **40**, 1843-1846.

129. M. T. Honaker, B. J. Sandefur, J. L. Hargett, A. L. McDaniel and R. N. Salvatore, *Tetrahedron Lett.*, 2003, **44**, 8373-8377.

130. E. A. Englund, H. N. Gopi and D. H. Appella, *Org. Lett.*, 2004, **6**, 213-215.

131. S. Ram and L. D. Spicer, *Tetrahedron Lett.*, 1987, **28**, 515-516.

132. T. Bieg and W. Szeja, *Synthesis*, 1985, 76-77.

133. G. Brieger and T. J. Nestrick, *Chem. Rev*, 1973, **74**, 567-580.

134. H. Yagi, D. R. Thakker, R. E. Lehr and D. M. Jerina, *J. Org. Chem.*, 1979, **44**, 3442-3444.

135. A. M. Felix, E. P. Heimer, T. J. Lambros, C. Tzougraki and J. Meienhofer, *J. Org. Chem.*, 1978, **43**, 4194-4196.

136. M. K. Anwer, S. A. Khan and K. M. Sivanandaiah, *Synthesis*, 1978, 751-752.

137. C. W. Alexander and D. C. Liotta, *Tetrahedron Lett.*, 1996, **37**, 1961-1964.

138. S. Ram and R. E. Ehrenkaufer, *Synthesis*, 1988, 91-95.

139. P. J. Finn, N. J. Gibson, R. Fallon, A. Hamilton and T. Brown, *Nucleic Acids Res.*, 1996, **24**, 3357-3363.

140. L. Kosynkina, W. Wang and T. C. Liang, *Tetrahedron Lett.*, 1994, **35**, 5173-5176.

141. E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem. Int. Ed.*, 1998, **37**, 2797-2823.

142. B. Hyrup and P. E. Nielsen, *Bioorg. Med. Chem.*, 1996, **4**, 5-23.

143. F. Albericio, J. M. Bofill, A. El-Faham and S. A. Kates, *J. Org. Chem.*, 1998, **63**, 9678-9683.

144. L. A. Carpino, *J. Am. Chem. Soc.*, 1993, **115**, 4397-4398.

145. R. F. Poulain, A. L. Tartar and B. P. Deprez, *Tetrahedron Lett.*, 2001, **42**, 1495-1498.

146. R. Knorr, A. Trzeciak, W. Bannwrth and D. Gillessen, *Tetrahedron Lett.*, 1989, **30**, 1927-1930.

147. J. J. Dudash, J. Jiang, S. C. Mayer and M. M. Jouille, *Synth. Commun.*, 1993, **23**, 349-356.

148. R. A. Torres and T. C. Bruice, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 649-653.

149. S. M. Chen, V. Mohan, J. S. Kiely, M. C. Griffith and R. H. Griffey, *Tetrahedron Lett.*, 1994, **35**, 5105-5108.

150. S. C. Brown, S. A. Thomson, J. M. Veal and D. G. Davis, *Science*, 1994, **265**, 777-780.

151. R. Alvarez, S. Velazquez, A. San-Felix, S. Aquro, E. De Clercq, C.-F. Perno, A. Karlsson, J. Balzarini and M. J. Camarasa, *J. Med. Chem.*, 1994, **37**, 4185-4194.

152. M. J. Genin, D. A. Allwine, D. J. Anderson, M. R. Barbachyn, D. E. Emmert, S. A. Garmon, D. R. Graber, K. C. Grega, J. B. Hester, D. K. Hutchinson, J. Morris, R. J. Reischer, C. W. Ford, G. E. WZurenko, J. C. Hamel, R. D. Schaadt, D. Stapert and B. H. Yagi, *J. Med. Chem.*, 2000, **43**, 953-970.

153. D. R. Buckle, C. J. Rockell, H. Smith and B. A. Spicer, *J. Med. Chem.*, 1986, **29**, 2262-2267.

154. A. C. Cunha, J. M. Figueiredo, J. L. M. Tributino, A. L. P. Miranda, H. C. Castro, R. B. Zingali, C. A. M. Fraga, M. C. B. V. de Souza, V. F. Ferreira and E. J. Barreiro, *Bioorg. Med. Chem.*, 2003, **11**, 2051-2059.

155. K. Dabak, Ö. Sezer, A. Akar and O. Anaç, *Eur. J. Med. Chem.*, 2003, **38**, 215-218.

156. V. D. Bock, H. Hiemstra and J. H. van Maarseveen, *Eur. J. Org. Chem.*, 2006, 51-68.

157. R. Huisgen, in *1,3-Dipolar cycloaddition chemistry, vol.1*, ed. A. Padwa, Wiley, New York, Editon edn., 1984, vol. 1, pp. 1-176.

158. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2002, **41**, 2596-2599.

159. C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057-3064.

160. H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004-2021.

161. F. Himo, T. Lovell, R. Hilgraf, V. V. Rostovtsev, L. Noodelman, B. K. Sharpless and V. V. Fokin, *J. Am. Chem. Soc.*, 2005, **127**, 210-216.

162. A. M. Jawalekar, N. Meeuwenoord, J. S. G. O. Cremers, H. S. Overkleeft, G. A. van der Marel, F. P. J. T. Rutjes and F. L. van Delft, *J. Org. Chem.*, 2008, **73**, 287-290.

163. W. H. Binder and C. Kluger, *Current Org. Chem.*, 2006, **10**, 1791-1815.

164. S. Y. Lim, W.-y. Chung, H. K. Lee, M. S. Park and H. G. Park, *Biochem. Biophys. Res. Commun.*, 2008, **376**, 633-636.

165. F. S. Hassane, B. Frisch and F. Schuber, *Bioconjugate Chem.*, 2006, **17**, 849-854.

166. V. Aucagne, J. Bern, J. D. Crowley, S. M. Goldup, K. D. Hänni, D. A. Leigh, P. J. Lusby, V. E. Ronaldson, A. M. Z. Slawin, A. Viterisi and D. Barney Walker, *J. Am. Chem. Soc.*, 2007, **129**, 11950-11963.

167. S. G. Davies, S. W. Epstein, A. C. Garner, O. Ichibara and A. D. Smith, *Tetrahedron: Asymmetry*, 2002, **13**, 1555-1565.

168. J. C. S. da Costa, K. C. Pais, E. L. Fernandes, P. S. M. de Oliveira, j. S. Mendonca, M. V. N. de Souza, M. A. Peralta and T. R. A. Vasconcelos, *Arkivoc*, 2006, **i**, 128-133.

169. J. Smith Gorsynski, S. E. Drozda, S. P. Petraglia, N. R. Quinn, E. M. Rice, B. S. Taylor and M. Viswanathan, *J. Org. Chem.*, 1984, **49**, 4112-4120.

170. M. C. Daga, M. Taddei and G. Varchi, *Tetrahedron Lett.*, 2001, **42**, 5191-5194.

171. R. W. Hoffmann and A. Hense, *Liebigs Ann.*, 1996, 1283-1288.

172. C. Flensbrug and M. Egholm, *Acta. Crystallogr. Sect. C*, 1994, **50**, 1480-1482.

173. B. E. Watkins, J. S. Kiely and H. Rapoport, *J. Am. Chem. Soc.*, 1982, **104**, 5702-5708.

174. B. E. Watkins and H. Rapoport, *J. Org. Chem.*, 1982, **47**, 4471-4477.

175. G. J. Sanjayan, V. R. Perdireddi and K. N. Ganesh, *Org. Lett.*, 2000, **2**, 2825-2828.

176. T. Govindaraju, R. G. Gonnade, M. M. Bhadbhade, V. A. Kumar and K. N. Ganesh, *Org. Lett.*, 2003, **5**, 3013-3016.

177. H. Ramussen, J. S. Kastrup, J. N. Nielsen, J. M. Nielsen and P. E. Nielsen, *Nature Struct. Biol.*, 1997, **4**, 98-101.

178. T. R. Chan, R. Hilgraf, B. K. Sharpless and V. V. Fokin, *Org. Lett.*, 2004, **6**, 2853-2855.

179. N. M. Howarth and L. P. G. Wakelin, *J. Org. Chem.*, 1997, **62**, 5441-5450.

180. C. Di Giorgio, S. Pairot, C. Schwergold, N. Patino, R. Condom, A. Farese-Di Giorgio and R. Guedj, *Tetrahedron*, 1999, **55**, 1937-1958.

181. N. M. Howarth, W. E. Lindsell, E. Murray and P. N. Preston, *Tetrahedron*, 2005, **61**, 8875-8887.

182. G. Procter, J. Leonard and B. Lygo, *Advanced Practical Organic Chemistry, 2nd Edition*, CRC Press, Cheltenham, UK, 1998.

183. H. E. Gottlieb, V. Kotlyar and A. Nudelman, *J. Org. Chem.*, 1997, **62**, 7512-7515.